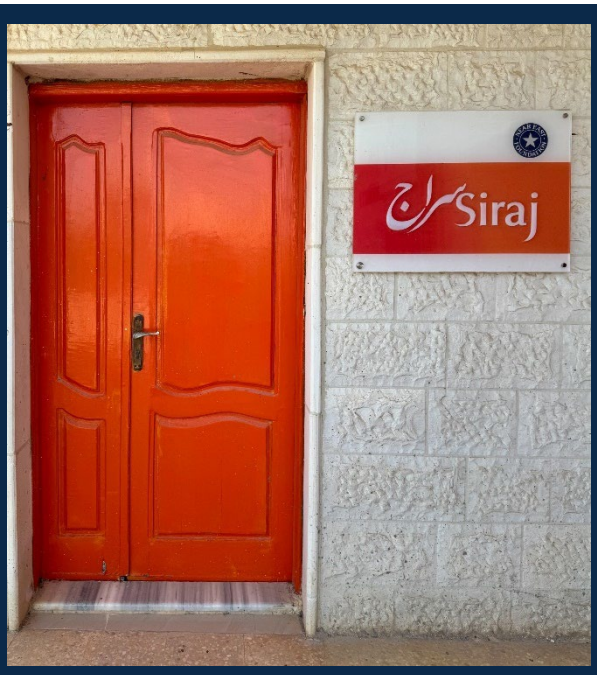
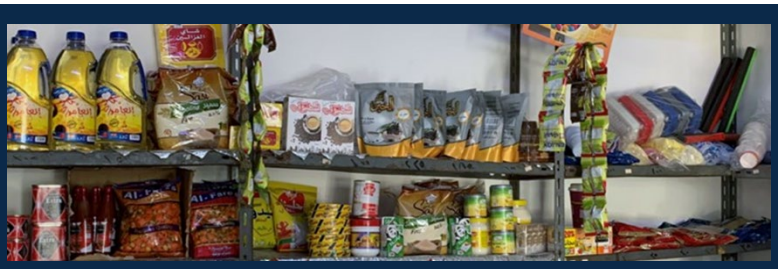
Mathematica
Progress Together





Evaluation of the Refugee Livelihoods Development Impact Bond in Jordan

Final report

September 2025

Laura Meyer, Evan Borkum, and Gray Collins

Submitted to:

Stichting IKEA Foundation
Dellaertweg 9G
2316 WZ Leiden, the Netherlands
Attention: Laura Costica

Submitted by:

Mathematica
1100 1st Street, NE, 12th Floor
Washington, DC 20002-4221
Phone: (202) 484-9220
Fax: (202) 863-1763

Contents

Acknowledgements

This report was made possible with the support of numerous people and organizations. First, we would like to acknowledge the guidance and support of the IKEA Foundation. We are especially grateful to Laura Costica and Hayley Kornblum for partnering with us closely throughout this process and providing valuable guidance on our approach and findings. We would also like to thank the other DIB partners and members of the steering committee, including KOIS, the Near East Foundation (NEF), the U.S. Development Finance Corporation. Ferd, the Novo Nordisk Foundation, and Norad.

Our thanks also go to NEF's program and monitoring and evaluation staff—especially Dalal Masalha, Sundus Al-Smadi, and James Murray—for their collaboration and transparency throughout the evaluation. They provided valuable insights into program implementation, willingly fielded our many requests for data, and successfully encouraged participation in the evaluation by program participants. Many other NEF program staff and their community-based organization partners generously shared additional insights on program implementation that were critical inputs to this report. We also thank Jonathan Peters for assistance collating the cost data for the cost effectiveness analysis.

We are also extremely grateful to our Jordanian evaluation partners, Mindset and Integrated International, for their support, expertise, and flexibility. We especially thank the Mindset team led by Majd Masannat for helping to ensure that we collected high-quality survey data that aligned with the evaluation timelines, even as those timelines shifted.

The evaluation team also benefitted from the contributions of other current and former colleagues at Mathematica. Jane Fortson, Nancy Murray, and Delia Welsh provided valuable feedback on draft memoranda and reports. Matt Spitzer contributed to the research design and early work on the evaluation. Paolo Abarcar supported the evaluation design and cost effectiveness analysis.

Finally, we wish to express our deep gratitude to the thousands of program participants who welcomed the evaluation team into their homes and shared their time and personal information about their daily lives. This evaluation would not have been possible without them, and we hope that in return the findings can inform efforts to improve the lives of the millions of others like them.

Executive Summary

Photo: The entrance to a Siraj Center hosted by a community-based organization where the program provided training and other services.

Introduction

The Syrian civil war, which began in 2011, generated the world's largest refugee crisis since the Second World War. As of 2024, Jordan hosted about 620,000 registered Syrian refugees, together with another 70,000 registered refugees from other crisis-hit countries (United Nations High Commissioner for Refugees 2024). A lack of livelihood opportunities forced many refugees to deplete their assets and savings, accumulate large debts, and resort to negative coping strategies. The influx of Syrian refugees occurred in a context in which there was a large population of vulnerable Jordanians, often in the communities hosting refugees, also facing livelihoods-related challenges. As the protracted nature of the displacement from Syria became apparent, the government of Jordan, foreign donors, and international organizations sought a long-term, development-oriented approach to build self-reliance and resilience among Syrian refugees. Aligned with this paradigm, a group of international organizations partnered on an innovative multi-year Refugee Livelihoods Development Impact Bond (DIB) in Jordan. This is one of 18 DIBs to date implemented in low- and middle-income countries, and the first one focused on improving the well-being of refugees through livelihoods programming. The DIB financed a microenterprise training and grants program for refugees and vulnerable Jordanians in host communities. The Near East Foundation UK (NEF) implemented the program in collaboration with local community-based organizations (CBOs). Under the DIB mechanism, DIB investors provided NEF with the upfront financing for the program and the DIB funders agreed to pay the investors at the end of the program based on the results achieved.

Key findings

- **NEF and their partner CBOs used data-driven adaptive management** to provide improved and more effective training and services to participants over time.
- **Grantees' businesses served as sustainable sources of income.** About three-quarters of grantees were still operating businesses after two years, generating average take-home business incomes of 98 Jordanian dinars (JOD) per month ($138 in nominal terms; $365 in purchasing power parity (PPP) terms) for household use.
- **Average annual household consumption was 636 JOD higher** for participants who completed the program two years prior than in a comparison group of future participants ($897, or PPP $2,366), driven by increased consumption of food and non-food goods and services.
- **More disadvantaged groups such as women, refugees, and poorer households experienced more barriers to entrepreneurship** and smaller impacts.
- **Impacts were almost exclusively driven by the receipt of cash grants**, with grantees experiencing an impact of 945 JOD ($1,332, or PPP $3,515) on annual consumption and non-grantees experiencing little impact.
- **The program model, including the CBO partnership approach, shows promise for adaptation and scaling to other contexts**, but there is room for further improvement, including through targeted supports to the most disadvantaged groups.

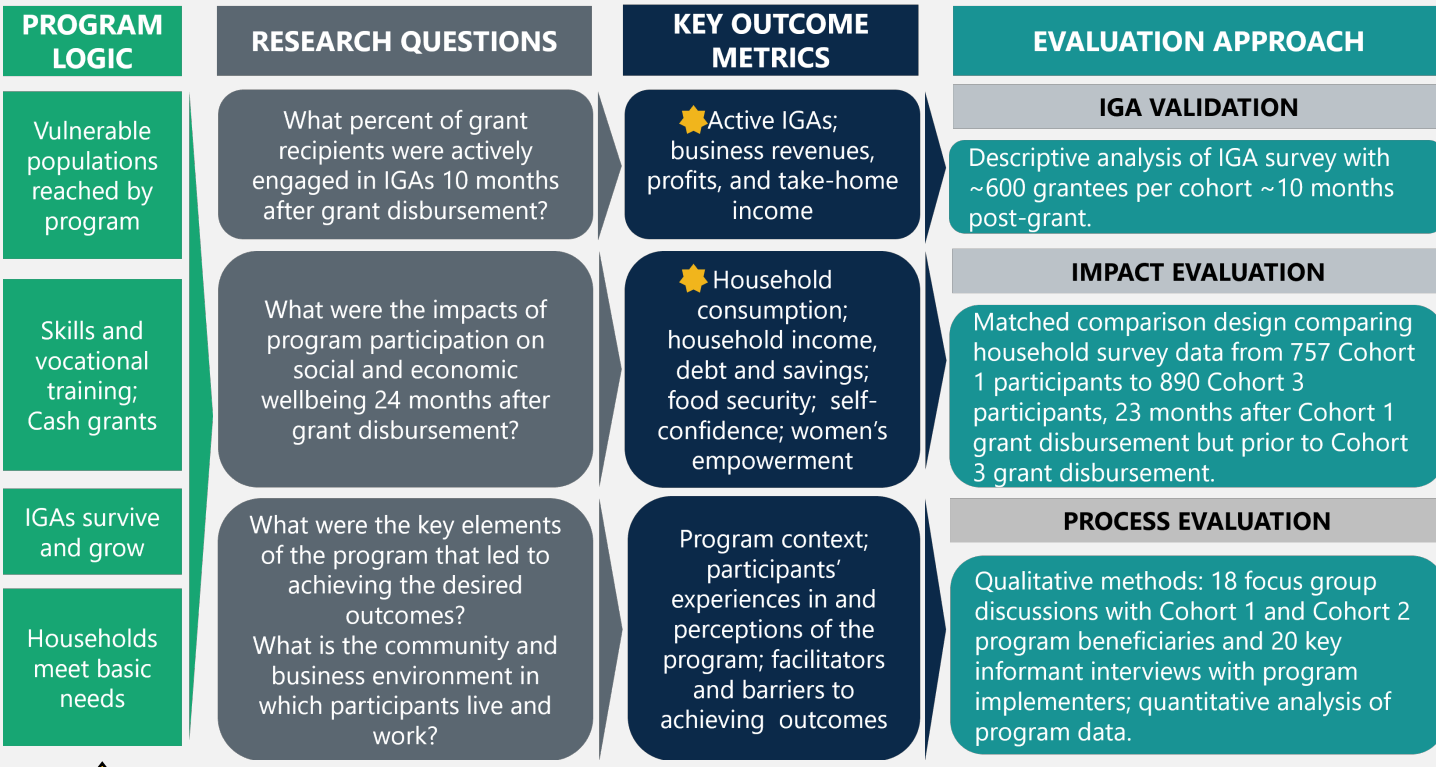
About the DIB program

The DIB-funded livelihoods program focused on supporting participants to create sustainable, mostly home-based, micro-enterprises. NEF partnered with local CBOs to identify participants based on a vulnerability assessment and deliver the program in five locations across Jordan. NEF and its partner CBOs served 5,660 participants across three program cohorts. More than three-quarters of participants were women, about one-third were refugees, about one-third were youth (ages between 18 and 25), and few were existing business owners. For each cohort, NEF and its CBO partners provided trainings and grants to small groups of participants over a six-month implementation period that started in April 2022 (Cohort 1), January 2023 (Cohort 2), or April 2024 (Cohort 3).

The core of the program was a five-day in-person sequenced training in business skills, culminating in the preparation of a business development plan that could potentially be funded through the program's cash grants. These grants were awarded to about 6 in 10 participants, ranged between 400 and 700 Jordanian dinars (JOD; between $564 and $987 in nominal terms, or between $1,488 and $2,604 in purchasing power parity [PPP] terms); the mean grant size was 580 JOD ($818, or PPP $2,158). Grant award was subject to an application from participants and approval of their plans by a grants committee. The program also provided additional post- grant support for grantees, primarily through one-to-one business mentorship sessions.

Mathematica conducted an independent evaluation of the program both to measure the metrics that determined payments to DIB investors and to generate insights about the program to support future adaptation and scale-up. The below figure summarizes the approach to the evaluation.

Evaluation framework and analytic approach



Notes: Metric triggering investor return (triggered above 44 percent for active IGAs and 0.22 standard deviations for consumption impacts); IGA = income-generating activity; ~ indicates an approximate or average number

Key findings

Program implementation

NEF's adaptive management approach led to continuous program improvement over time. In response to feedback from Cohort 1, the program made trainings more practical, improved flexibility and accessibility to accommodate participants' other life responsibilities, provided additional support to grantees in spending their grant, and improved communication during the grant selection process.

The core business skills training, cash grants, and post-grant support provided a strong foundation for participants' microenterprises. Participants perceived the business skills training as inclusive and highly valuable. The program used a rigorous grant selection process to identify proposed businesses with strong potential for success. It was common for grantees to invest the entire grant in their businesses, and many supplemented it with personal resources. During the post-grant period, grantees benefited substantially from one-to-one mentorship, during which trainers conducted site visits to grantees to provide refresher trainings and offer support and encouragement.

However, there is still room for future improvement to facilitate applicants' success in applying for and using grants. The interview that was part of the grant selection process was anxiety-inducing and uncomfortable for many participants, especially women. Future iterations of the program could consider offering more details about the selection criteria, providing additional interview preparation, or taking other steps to mitigate the anxiety around the grant selection process. Further, the grant ceiling posed a constraint to start-up and growth for some grantees whose businesses were capital-intensive, based outside the home, and/or operating in Amman, where costs tended to be higher.

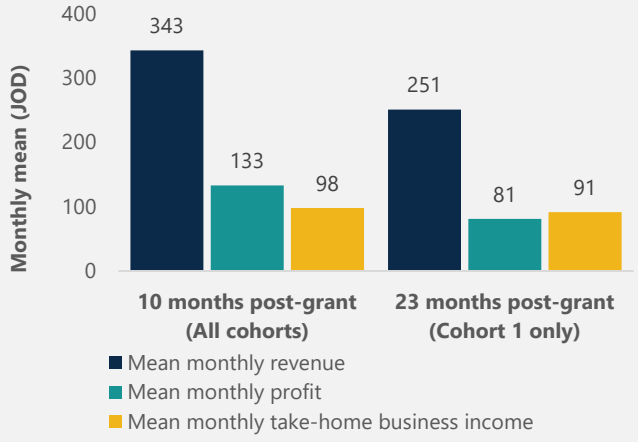
Photo: A participant who launched a bakery with support of the project shows off a cake they produced. The top of the cake bears NEF's logo. Home food processing was the most common type of business, accounting for about one quarter of businesses supported by the program (according to the IGA survey), and was even more common among women.

Program effects on income-generating activities

The business skills training helped participants develop critical skills to successfully establish and operate their businesses. Grantees put these skills into practice, and most reported implementing small business management best practices that are typically associated with other positive business outcomes. Participants also reported a greater sense of self-confidence, motivation, and independence as a result of the training.

About 10 months after grants were disbursed, almost all grantees' businesses were still active, and the vast majority were earning positive monthly profits and generating income to support personal and household expenses. The main payment metric for the DIB, the business metric defined based on having an active IGA 10 months after grants were disbursed, was 98 percent for grantees across all cohorts combined. These active businesses were typically conducting frequent transactions and almost 90 percent were earning positive profits. Mean monthly profits were 133 JOD ($188, or PPP $495), of which a mean of 98 JOD ($138, or PPP $365) was take-home income that went towards supporting personal and household expenses.

Business financial metrics for active grant-supported businesses at 10 and 23 months post-grant



Source: IGA surveys (10 months) and impact survey (23 months)

Mean reported take-home income is equivalent to about one-third of the national minimum wage and about one-third of mean monthly household expenditures for refugees. Male grantees reported higher levels of revenue, profits, and take-home income from businesses than female grantees. This is likely related to differences in business types, the additional resources they have invested in their businesses, and the amount of time they spend each week on their businesses.

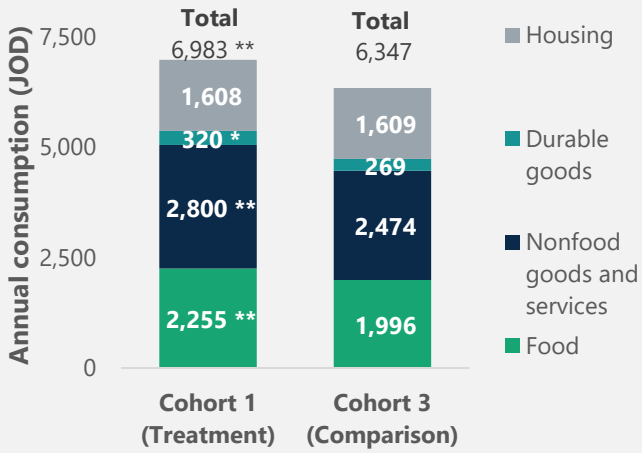
Almost two years after grant disbursement, most grant-supported IGAs from Cohort 1 were still active and serving as a steady source of income. In the two-year impact survey, 76 percent of Cohort 1 grantees still satisfied the criteria used to define an active IGA for the DIB's 10-month business metric. Most Cohort 1 grantees' businesses remained profitable about two-years after receiving grants. Mean reported revenues and profits among active Cohort 1 businesses declined relative to the 10-month mark, although we cannot rule out that this is related to poorer business record-keeping over time, with more self-reports at the two-year mark. Nevertheless, mean take-home business income for personal and household expenses among active businesses, which was self-reported at both 10 months and two years, held steady over time .

Program effects on well-being

The program led to a 10 percent, or 0.22 standard deviation increase in total annual household consumption; a separate analysis showed that impacts were predominantly experienced by grantees. Almost two years after the grants were disbursed to Cohort 1**,** the estimated mean value of the household consumption metric for Cohort 1—including grantees and non-grantees—was 636 JOD ($897, or PPP $2,366) higher than matched Cohort 3 households. This impact was driven almost entirely by Cohort 1 grantees, who experienced an impact of 945 JOD ($1,332, or PPP $3,515), equivalent to a 15 percent or 0.36 standard deviation increase relative to matched Cohort 3 future grantees. In contrast, there were near-zero impacts on non-grantees. This implies that receipt of grants and post-grant support, rather than the business development training, are driving the overall impacts on consumption .

Households were using most of their increased income to increase consumption of nutritious and staple foods, increase their use of health care services, and meet other basic needs like clothing and utilities. Most of the impacts on consumption were driven by increased consumption of food and non-food goods and services.

Impacts on annual household consumption, overall and by category



Source: Impact survey

Notes: Samples include grantees and non-grantees.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test

Households also spent their increased income on increasing their household assets, primarily household appliances and electronics. Cohort 1 households reported modestly reduced food insecurity and utilization of harmful coping strategies compared to matched Cohort 3 households. Despite these positive impacts, it was still common for Cohort 1 households to use relatively severe coping strategies, suggesting that most were still not able to ensure food security and fully meet their basic needs.

Impacts on income and consumption were not evenly distributed across different sub-populations, and tended to favor groups who were more advantaged prior to starting the program. Men, youth, Jordanians, and households that were relatively better off prior to joining the program experienced the largest impacts on income and/or consumption. However, differences in impacts across subgroups are complex, and likely reflect an interplay of demographic, socio-economic, and other household characteristics, as well as unmeasured expenditure categories like debt repayments and remittances.

Conclusions and recommendations

The impacts on the consumption metric are near the upper range of impacts found in the reference studies that were used to set the thresholds for DIB payments. Standard deviations were used as the unit of the consumption metric because they are a common way to compare impacts across different outcomes and program contexts. The reference studies had impacts of between 0.07 and 0.38 standard deviations on consumption or expenditure, a closely related measure. In comparison, our estimated impacts on consumption were 0.22 standard deviations for all participants, and 0.36 standard deviations for grantees only. Impacts were also similar to impacts of livelihoods and cash transfer treatments from the Middle East and North Africa (MENA) region.

Recommendations for livelihoods program

- ✓ **Carefully select and build the capacity of CBOs** to serve as vital partners throughout implementation.
- ✓ **Include cash grants or tailored in-kind support** in entrepreneurship programming to help participants overcome financial constraints.
- ✓ **Carefully identify participants with the motivation and ideas to be entrepreneurs** but who may lack sufficient resources or skills to launch or grow businesses.
- ✓ **Provide additional, targeted supports** to subgroups who face barriers to income-generating activities.
- ✓ **Consider results-based funding models** that provide stable funding and flexibility for implementer-led innovation.

Recommendations for results-based financing programs

- ✓ **Align payment metrics with a detailed program logic**, including both short-term and long term outcomes.
- ✓ **Consider using household expenditures, rather than household consumption,** as a practical measure of economic well-being.

The multi-year flexible funding provided by the DIB, its use of both short- and longer-term payment metrics, and multiple stages of measurement, helped to align implementer incentives with program objectives and support program improvements over time. The guaranteed funding and programmatic and budgetary flexibility offered by the DIB funding model has encouraged NEF to test multiple activities and approaches, collect and analyze data at each phase to reflect on their effectiveness, and improve their approaches over time. Internal and external monitoring and evaluation activities have resulted in measurable improvements in implementation across cohorts. The DIB payment metrics also struck a good balance between balancing DIB parties' financial risk and sufficiently incentivizing sustainable improvements in outcomes. Further, the multi-cohort approach and multi-step evaluation has fostered a collaborative, mutually supportive relationship between NEF and the independent evaluation team, supporting ongoing improvements in program implementation and data quality.

The program was cost-effective. We estimate that, over 10 years, the program will generate net benefits of about $2,900 per grantee in business profits and $3,500 per grantee in household consumption after subtracting program costs. The benefit-cost ratio for business profits is 1.98, meaning the program generated $1.98 in profits for every dollar invested. For household consumption it is 2.16, which compares favorably to related programs.

Expenditures may be more suitable for use as a measure of economic well-being and a DIB payment metric in this context than consumption. Despite its theoretical advantages, measuring consumption posed some challenges for survey respondents and omitted expenditure categories that reflect household economic well-being and may be important in this context. Although measuring expenditures also has some disadvantages, an expenditure-based measure might have been preferable given that the aim of the evaluation was to compare economic well-being between a treatment and comparison group rather than to produce an accurate stand-alone measure of household consumption.

The use of local CBOs as a hub for services can strengthen implementation effectiveness and sustainability. Interviews with program implementers indicated that CBOs played a critical role in the success of implementation, using their longstanding presence in the community to support broad-based recruitment efforts, build participant trust and confidence in the program, and address key barriers to participation. NEF also reported engaging CBO staff in the program design stage, collecting participant feedback, assessing implementation strategies, reviewing, and interpreting monitoring data, and informing adaptations and improvements to activities across cohorts. In turn, collaborating with NEF strengthened CBO capacity to implement similar programs, resulting in increased outside funding. This program highlights the value of locally led implementation of livelihood programs, with appropriate support and capacity building from larger national or international organizations with the relevant capacity, experience, and local knowledge.

The positive findings suggest that the program was effective for participants selected for grants, but it may not be a catch-all solution for improving the well-being of all vulnerable populations. These results reflect benefits for a carefully selected group of vulnerable individuals who demonstrated the aspirations and the capacity to be entrepreneurs. It is unlikely that the program would be similarly effective if it were scaled up in a way that involved a less stringent selection process that sought to reach a broader vulnerable population. The program also relied on NEF's extensive experience with and learnings from implementing similar programs in the MENA region and its deep understanding of the cultural context. Adapting this program to other countries or by other implementers would need to carefully account for the local business environment and economy; social, cultural, and gender norms; and implementer experience.

Photo: The entrance to a CBO center. Partnering with CBOs helped NEF recruit participants and adapt the program to the needs of local communities.

While the findings overall are positive, they also suggest that additional, targeted supports may be needed to ensure that the benefits of the program are distributed more equitably. Subgroup findings show that some subgroups (women, refugees, the economically worst-off at baseline) experienced smaller impacts on income and/or consumption than others, and qualitative evidence suggests that women and refugees faced additional barriers to starting and growing their businesses. A comprehensive assessment to identify the primary barriers for these groups and targeted supports that could address those barriers in the local context could help to promote greater equity in program outcomes.

I. Introduction

The Syrian civil war, which began in 2011, generated the world's largest refugee crisis since the Second World War. Almost 6.6 million people—about one quarter of Syria's pre-war population—are estimated to have fled the country, with most seeking refuge in neighboring Turkey, Lebanon, and Jordan. As of late 2024, Jordan hosted about 620,000 registered Syrian refugees, together with another 70,000 registered refugees from other crisis-hit countries such as Iraq, Yemen, and Sudan (United Nations High Commissioner for Refugees 2024).

Among Syrian refugees in Jordan, about 8 in 10 lived below the national poverty line in 2019, even before the COVID-19 pandemic worsened economic conditions in the country (United Nations High Commissioner for Refugees 2019). Female-headed refugee households are particularly vulnerable, with cultural expectations and home care duties acting as additional constraints to securing employment and livelihoods. An estimated 40 percent of refugee households are headed by women and, as of 2016, only 3 percent of refugee women worked (Krafft et al. 2019; United Nations 2018). Although Jordan has several refugee camps in which governmental and non-governmental organizations provide free services, limited livelihood opportunities and a lack of privacy in the camps lead most refugees to live in cities outside the camps, where they face high housing costs (Aziz et al. 2019; Wall et al. 2017; United Nations High Commissioner for Refugees 2019). A lack of livelihood opportunities has forced many refugees to deplete their assets and savings and accumulate large debts (Culbertson et al. 2016, ReliefWeb 2017). Many refugees have also resorted to other negative coping strategies such as accepting socially degrading, exploitative, high risk, or illegal jobs or reducing expenditures on essentials (United Nations High Commissioner for Refugees 2019).

The large influx of Syrian refugees occurred in a context in which there is a large population of vulnerable Jordanians—often living in the same communities as refugees—facing livelihoods-related challenges. For example, in 2018, 19 percent of Jordanian adults were unemployed and about 16 percent of Jordanians lived below the poverty line (United Nations Children's Fund 2020). The influx of refugees has affected Jordanian citizens in complex ways, both real and perceived. For example, increased competition for housing may have worsened average housing quality for poor Jordanians and increased rental prices (Al-Hawarin et al. 2018). In contrast, recent evidence suggests that the labor market outcomes of Jordanian workers were little affected, likely because refugees and Jordanians were largely not competing for the same jobs and because increased public sector investment in response to the refugee crisis improved job opportunities for Jordanians (Fallah et al. 2018; Malaeb and Wahba 2018). Nevertheless, resentment over differential access to employment and perceptions of Syrians competing with Jordanians for jobs might still have increased tensions and adversely affected social cohesion in host communities (REACH 2014).

In the first years of the Syrian refugee crisis, the response by the government of Jordan, foreign donors, and international organizations was primally humanitarian in nature, focused on meeting refugees' short-term needs for shelter, food, and cash. However, as the protracted nature of the displacement from Syria became apparent, these stakeholders sought a longer-term, more development-oriented approach to build self-reliance and resilience among Syrian refugees. Such an approach, which focuses on helping refugees secure sustainable livelihoods, could position refugees to become assets and major economic contributors to their host countries in the long run (Clemens et al. 2018; Legrain and Burridge 2019).

An important step in this new approach was the 2016 Jordan Compact, signed by the Jordanian government and the European Union, which outlined the objectives of facilitating refugees' labor market access while mitigating adverse impacts to local citizens through financial assistance and trade concessions (Lenner and Turner 2019). Under the Compact, the Jordanian government agreed to allow as many as 200,000 Syrian refugees to obtain work permits in certain sectors, simplified the fees and administrative procedures for these permits, and allowed Syrians to operate certain types of home-based businesses. This initiative is believed to have contributed to a modest increase in labor force participation among adult Syrian refugees and a large decrease in the unemployment rate among adult Syrian refugees in the labor force between 2014 and 2018 (Tiltnes et al. 2019). Nevertheless, in 2018, about 40 percent of adult Syrian refugees remained out of the labor force and a further 15 percent were in the labor force but unemployed. Among those employed, the majority were employed in the informal sector, where their jobs lack legal protection. This suggests that, despite improvements by the start of the 2020s, many refugees were still unable to attain secure livelihoods.

Building on the paradigm shift toward a development-oriented approach to the Syrian refugee crisis, a group of international partners collaborated on an innovative Refugee Livelihoods Development Impact Bond (DIB) in Jordan. The DIB, coordinated by KOIS, financed a four-year microenterprise training and grants program for refugees and vulnerable Jordanians in host communities. This program was implemented by the Near East Foundation UK (NEF) in collaboration with local community-based organizations (CBOs). The DIB investors (the United States International Development Finance Corporation [DFC] and Ferd, a family-owned Norwegian investment company) provided NEF with the upfront financing for the program. Under the DIB mechanism, the funders (IKEA Foundation, Novo Nordisk Foundation, and Norad) agreed to pay the investors at the end of the program, with the final payment amount depending on the results achieved. The main objectives of this DIB were: (1) to sustainably improve program participants' abilities to meet basic needs; their economic well-being, self-reliance, and resilience; and women's confidence, bargaining power, and agency; and (2) to encourage international development actors to devote more resources to long-term livelihoods programs in refugee contexts and demonstrate the potential of innovative funding mechanisms to achieve this.

The Refugee Livelihoods DIB is one of 18 DIBs in low- and middle-income countries launched since 2018 (Brookings Institution, 2025).[1] DIBs have become an increasingly popular mechanism for funding development programs and offer two main benefits compared to traditional grant-based financing. First, they appeal to a wider range of potential investors, including those in the private sector, who have access to large pools of capital. This could potentially lead to larger volumes of capital being mobilized for development programs and thus increased programmatic scale. Second, they intensify all partners' focus on measurement, learning, and results, given the payments at stake. This has the potential to lead to improved quality in program design and implementation and thus increased effectiveness of the program. Whether, and to what extent, a DIB-funded program in fact achieves results that differ from a traditional-grant funded one has not been rigorously measured, and likely depends on the context. Previous DIBs in low- and middle-income countries have tackled challenges in health (Cameroon, India, Nigeria,

[1] There have also been 16 Social Impact Bonds (SIBs) launched in low- and middle-income countries since 2018. Under a SIB, the outcome payer is the domestic government instead of a foreign government or private foundation, as is the case with a DIB.

Democratic Republic of Congo, Mali), employment and training (Palestine), education (India), poverty reduction (Kenya), and agriculture and the environment (Peru) (Carter et al. 2024). To the best of our knowledge, the Refugee Livelihoods DIB is the first to focus on supporting livelihoods in a refugee context outside of a handful supporting refugee integration in high-income countries.

The IKEA Foundation contracted with Mathematica to conduct an independent evaluation of the DIB program. The evaluation sought to both measure the metrics to determine payments to investors and generate broader learning about the program's impacts to support future adaptation and scale-up. In this final evaluation report, we begin by providing a brief overview of the DIB program in Section II, including the program activities, DIB structure, and program logic. In Section III, we describe our mixed-methods methodological approach to the evaluation. In subsequent sections, we present the evaluation findings, which we organize based on the program logic into findings on implementation and immediate program outputs (Section IV), short-term outcomes (Section V), and medium- and long-term outcomes (Section VI). We conclude in Section VII by summarizing the findings and lessons learned, including a discussion of how findings compare to evidence on the impacts of other, similar programs.

II. The Refugee Livelihoods DIB Program

In this section we describe the DIB program and the structure of the DIB itself. We also discuss how the program activities were expected to lead to the desired outcomes according to the program logic, which we use to organize the findings in subsequent sections.

A. Program description

The DIB-funded livelihoods program focused on supporting participants to create sustainable, mostly home-based, micro-enterprises. The program provided short trainings in business skills, life skills, and technical/vocational skills (for some participants); cash grants to finance micro-enterprises for participants' with sufficiently strong business plans; and additional technical and financial support to participants selected for grants. It built on similar work that NEF has conducted in Jordan and Lebanon since 2013 through several iterations of the Enhancing Economic Resilience project.

NEF partnered with local CBOs, which were selected on a competitive basis, to identify participants and deliver the program in five locations across Jordan: Amman, Irbid, Kufrsoum, Russeifa, and Zarqa. In each location. NEF and its partner CBO delivered the program to three program cohorts through NEF's existing Siraj centers. Siraj centers are physical hubs managed by the CBOs at which vulnerable individuals can access training, financial resources, and advisory services to support their livelihoods, or find referrals to or information about other services related to their physical, economic, and mental wellbeing. By training and supporting local CBOs to implement the program, NEF hoped to build their capacity to identify and support refugees and vulnerable Jordanians.

To recruit participants for each of the three program cohort, NEF and its partner CBOs conducted door-to-door canvassing, coordinated referrals (for example, from municipalities, other CBOs, non-governmental organizations, United Nations agencies, or government anti-poverty programs), and welcomed Siraj center walk-ins. Potential participants were screened using a vulnerability assessment tool to identify those who most needed livelihoods support to meet their basic needs, were interested in starting or growing a business, and were willing to commit to the program. Ultimately, NEF and its partner CBOs served 5,660 participants across the three program cohorts (**Table II.1)**. The cohorts were similar in their socio-demographic composition. More than three-quarters of participants were women, about one-third were refugees, about one-third were youth (ages between 18 and 25), and few were existing business owners.

Table II.1. Participants served by the Refugee Livelihoods DIB Program

	Cohort 1 (started April 2022)	Cohort 2 (started January 2023)	Cohort 3 (started April 2024)	Total
Participants served	1,235	1,902	2,523	5,660
Women	82.4%	81.9%	79.6%	81.0%
Refugees	30.4%	30.2%	34.0%	31.9%
Youth	30.1%	31.5%	30.9%	31.0%
Had existing business	10.4%	9.0%	6.2%	8.1%

Source: NEF program activity data

The program was designed to build participants' skills and support the development and growth of their businesses through the following activities (**Table II.2** shows the percentage of participants who engaged in each activity):

1. A core five-day in-person **sequenced training in business skills** for all participants, culminating in the preparation of a business development plan that could potentially be funded through the program's cash grants.
2. A two-day virtual **training in life skills**, which was also intended to be a core training for all participants. This was ultimately offered only to grantees in Cohort 1 because of implementation delays but was offered to all participants in subsequent cohorts. NEF reported that in cohorts 2 and 3, participation rates among grantees was high (more than 95 percent) but low among non-grantees, leading to differences in participation across cohorts that reflects the relative proportion of grantees (see below).

Table II.2. Program engagement

	Cohort 1	Cohort 2	Cohort 3
All participants			
Business skills training	100%	100%	100%
Life skills training	66%	89%	49%
Technical/vocational skills training	25%	35%	26%
Cash grants	67%	63%	55%
Grantees only			
One-to-one mentorship	99%	98%	96%
Small group coaching	89%	92%	92%
Advanced grants	6%	0%	0%
Advanced training	47%	54%	0%

Source: NEF program activity data

Notes: The total number of participants varied by cohort: 1,235 for Cohort 1, 1,902 for Cohort 2, and 2,523 for Cohort 3.

3. **Technical/vocational skills training** lasting between 3 and 6 days (depending on the topic), for about one-quarter of Cohorts 1 and 3 and one-third of Cohort 2 participants in topics relevant to their business development plan, conducted by third party experts or training providers.
4. **Cash grants**, awarded to about 6 in 10 of all participants to finance their business development plans. Rates of grant receipt were lowest in Cohort 3 (55 percent of participants): due to widespread interest in the program, NEF increased the pool of trainees, but was unable to proportionally increase the number of available grants. Grants ranged between 400 and 700 Jordanian dinars (JOD; between $564 and $987 in nominal terms, or between $1,488 and $2,604 in purchasing power parity [PPP] terms).[2] Across all three cohorts, the mean grant size was 588 JOD ($829, or PPP $2,187). Grant award was subject to an application from participants and approval of their plans by a grants committee. Jordanian applicants were required to commit to spending the grant on business-related needs; refugees were not required to make commitments about how the grant was to be used.
5. Additional advanced and specialized support for grantees as part of the "Siraj accelerator" initiative, driven by grantee demand and informed by a monthly business support tool that tracked the

[2] The nominal exchange rate is pegged, and JOD can be converted to dollars by multiplying by 1.41. One can also convert JOD to dollars using a PPP-adjusted exchange rate, which accounts for differences in the cost of living, by multiplying JOD by 3.72 (International Monetary Fund, 2024).

challenges and support needs of each grantee. This support included **one-to-one business mentorship sessions and small group coaching**. On average, grantees participated in between 4 and 5 one-on-one mentorship sessions provided by trainers and 1 small group coaching session.

6. An **advanced tailored business training for members of Cohorts 1 and 2** based on their individual needs. Although advanced training was not provided for Cohort 3 in the same format as for Cohorts 1 and 2, aspects of this training were folded into the small group coaching curricula.
7. **Additional financial support for selected grantees** whose businesses demonstrated particularly strong potential for growth. For Cohort 1, about 6 percent of grantees were selected through a competitive process to receive advanced grants, with amounts of between 1,500 and 2,000 JOD (between $2,115 and $2,820, or between PPP $5,580 and $7,440). The selection process was similar to that used to select grantees for the initial grants. For Cohorts 2 and 3, resources for advanced grants were reallocated to different program activities, and instead all grantees (including those from Cohort 1) were encouraged and supported to apply for loans through banks, microfinance institutions, and a NEF-financed revolving fund mechanism that preceded and is not directly related to the DIB. NEF reports that 69 Jordanian grantees from Cohort 1 and 2 have received loans of approximately 1,000 JOD ($1,400, or PPP $3,720) each. As of the time this report was published, no Cohort 3 grantees have received loans from this fund.

For each cohort, NEF and its CBO partners provided the core business and life skills trainings, technical/vocational skills training, and grants to participants over a six-month implementation period. NEF staff oversaw curriculum development, selected and trained trainers, and managed the implementation of training, grantee selection, and post-grant-supports with continued input from and collaboration from its CBO partners. Participants in each cohort were divided into small groups that received training at different times within the implementation period; grants were awarded in the second half of the implementation period, based on the assessment of a business plan developed during the core training. NEF and its partners provided Siraj accelerator support, advanced business training, and other ongoing support in subsequent months.

Over the course of the program, NEF used an adaptive management program implementation approach that drew on monitoring data and other feedback from implementers and participants to continually adapt implementation. As a result, the types and duration of training and supports varied slightly across cohorts. Section IV summarizes findings on program implementation and the changes that implementers and participants reported over time.

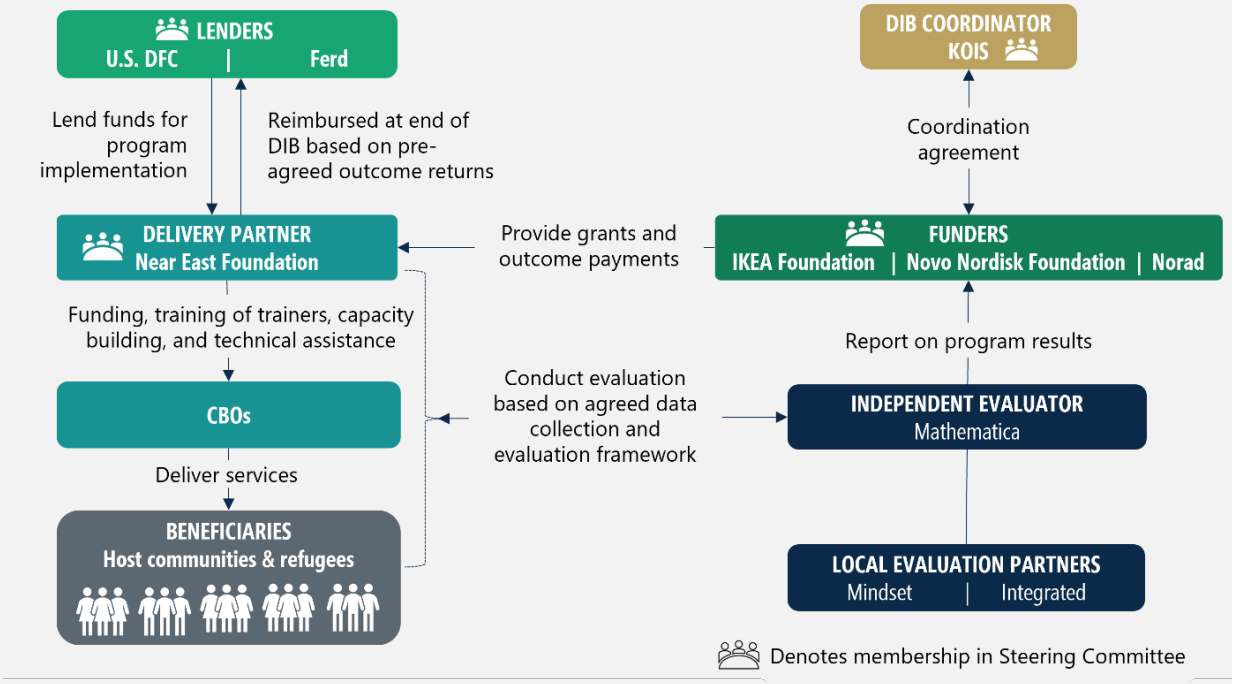
B. Structure of the DIB

The four-year, $9.8 million DIB included several parties, each with a distinct role (**Figure II.1**). NEF (the service provider) received up-front capital from DFC and Ferd (the lenders) to implement the program for each cohort, working through the local CBOs. Mathematica —in close collaboration with its local evaluation partners, Mindset and Integrated International conducted an evaluation of the DIB, including measuring the two DIB payment metrics: (1) the percent of grantees across all three cohorts actively engaged in IGAs about 10 months after grants were disbursed (the "business metric"), and (2) impacts on household consumption for the first cohort almost 24 months after grants were disbursed (the "household consumption metric" or "consumption metric"). The up-front capital committed by the lenders

fully covered NEF's program implementation costs. At the conclusion of the DIB, the IKEA Foundation and Novo Nordisk Foundation (the outcome funders) agreed to make payments to NEF to enable them to repay the lenders a minimum of 80 percent of this up-front capital and up to 122 percent of this amount (equivalent to a 5.1 percent annual rate of return). The total payment was dependent on the value of the two payment metrics reported by Mathematica, with the minimum payment thresholds for the business metric and the consumption metric set at 44 percent and 0.22 standard deviations, respectively.[3] Norad (an additional funder), also made payments to NEF through a grant to reimburse program implementation costs funded by the 80 percent capital guarantee.

Under this DIB model, the lenders take on the risk of program performance, the service provider benefits from the stability of multiple years of funding, and the outcome funders repay lenders based on concrete achievements in targeted social outcomes that are verified by the independent evaluator. KOIS (the DIB coordinator) is responsible for the financial structure of the DIB, coordination of agreements between the other parties, engagement of service providers, and coordination between parties during implementation.

Figure II.1. Structure of the DIB



[3] Beyond the base payment, which is equivalent to 80 percent of the up-front capital committed, investors receive an additional payment for each tenth of a percentage point by which the business metric exceeds 44 percent, up to a cap of 75 percent. The payment per tenth of a percentage point decreases once the business metric exceeds 55 percent, the minimum target for the DIB. DIB payments for the household consumption metric are only made if the business metric exceeds the minimum target of 55 percent, which is equivalent to the total loan repayment without interest. In that case, payments to investors begin for impacts on household consumption of 0.22 standard deviations and increase for each additional impact of 0.01 standard deviations up to a cap of 0.38 standard deviations. In addition, NEF is entitled for a success bonus which increases incrementally by a tenth of a percentage point for results between 67.6 percent and 75 percent for the business metric.

C. Program logic

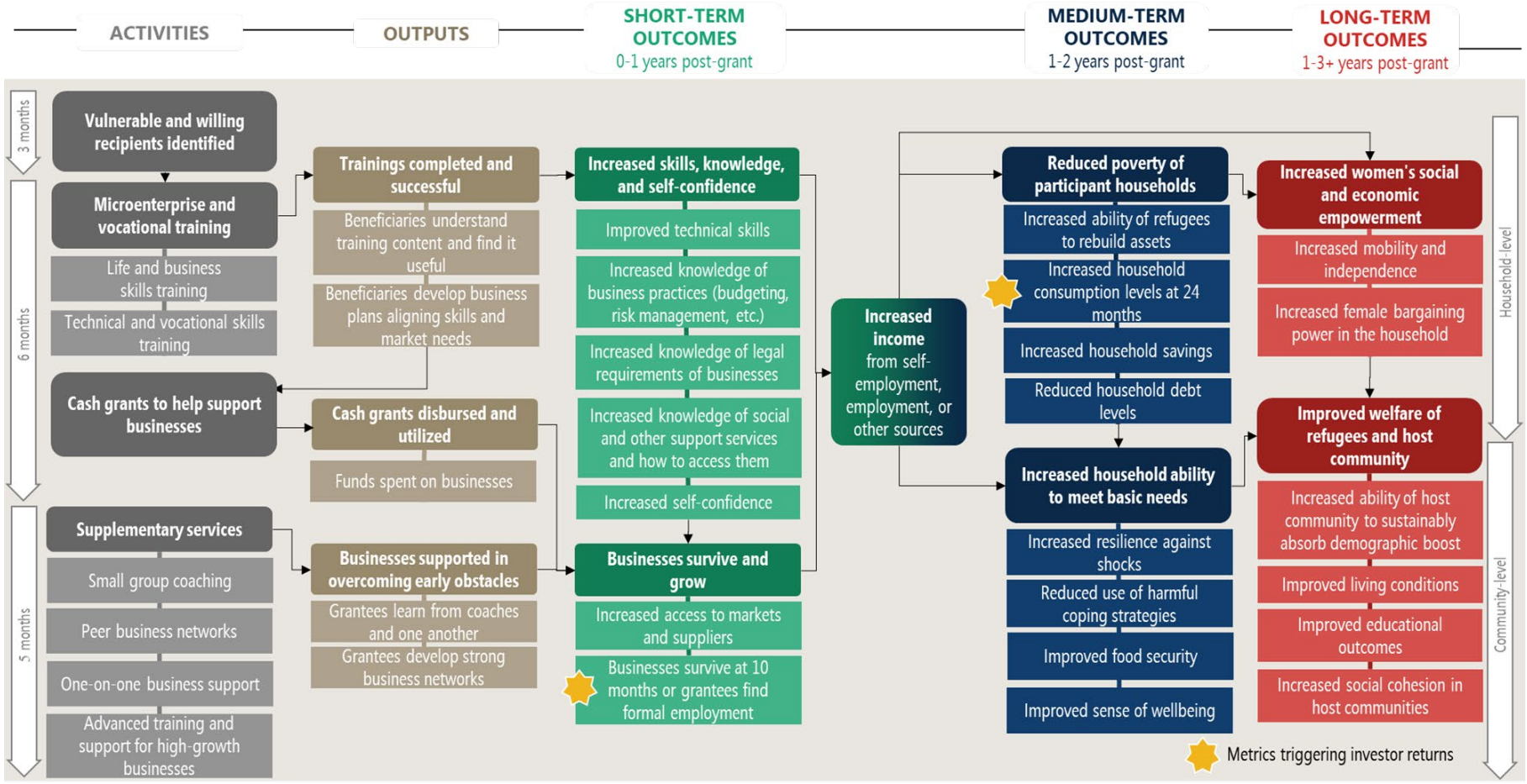
The program logic (**Figure II.2**) illustrates how the program's activities were expected to lead to outputs and subsequent outcomes. The training activities conducted during the implementation period were intended to culminate in participants developing viable business plans and cash grants being awarded to those with strong plans. A set of supplementary services was designed to support grantees to overcome early obstacles in establishing or expanding their businesses as income generating activities (IGAs).

In the short-term, program participants were expected to build their knowledge of business practices and legal requirements, technical skills, and self-confidence. The increase in knowledge and skills, together with the post-grant support, would then help these newly established businesses survive and grow. It was anticipated that some participants might use their new skills, self-confidence, and increased knowledge of other support services to find formal wage-earning employment as an alternative source of income.

In the medium term, these short-term outcomes were expected to translate to a sustained increase in participants' income from self-employment or wage employment. This increased income would reduce poverty as refugees rebuilt their assets, increased their savings, reduced their debt levels, and increased consumption. As a result, households would be better able to meet their basic needs without having to resort to harmful coping strategies, leading to a greater overall sense of wellbeing.

In the longer term, these changes were expected to contribute to broader transformations both at the household and at the community level. Female participants' increased contribution to household income would increase their social and economic empowerment within the household and the community over time. In the aggregate, host communities would be better able to integrate refugees as both refugees and vulnerable Jordanians were better able to meet their basic needs and invest in education and tangible assets, resulting in improved living conditions and social cohesion.

Figure II.2. Program logic



Notes: This figure differs slightly from that presented in the evaluation framework report. Increased household ability to meet basic needs is now a medium-term rather than a long-term outcome since it is a more immediate result of increased income. Conversely, increased women's social and economic empowerment is now a long-term outcome rather than a medium-term outcome because we anticipate more time may be required to change social and gender norms. These changes do not affect the evaluation design because we measured both medium- and long-term outcomes at the same time, two years post-grant. We have also added improved educational outcomes as a measure of improved welfare because we included educational outcomes in the evaluation framework based on an implicit connection between education and well-being.

III. Methodological approach

In this section we present the key research questions that the evaluation seeks to address and describe the evaluation design that we use to answer them. We provide additional methodological details in Appendices B, C, and D.

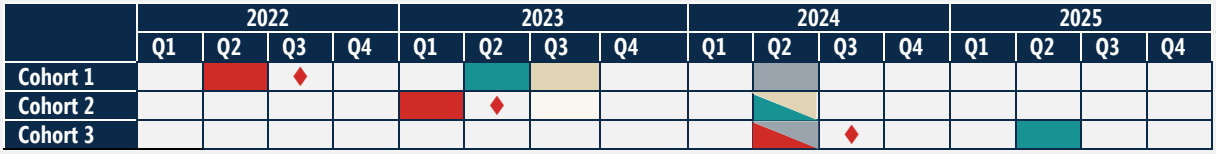
A. Research questions and overview of evaluation design

Table III.1 presents the research questions for the evaluation, which are underpinned by the program logic described in Section II. These questions cover the two DIB payment metrics (shown in bold), but also address effects of the program on participants' social and economic wellbeing, as well as the mechanisms and context underlying these effects. To answer these questions, we use a rigorous mixed-methods evaluation comprising three components (**Figure III.1** summarizes the timeline of key evaluation activities relative to program implementation):

1. **An IGA validation study** to measure the primary metric used to determine DIB payments, the business metric, which is defined as the percentage of grant recipients who were actively engaged in IGAs about 10 months after the midpoint of grant disbursement for each cohort. Data to assess performance on this metric comes from a short survey with a representative sample of grant recipients from each of the three cohorts ("the IGA survey").
2. **An impact evaluation** to assess the impacts of the program on household consumption, the secondary DIB payment metric, and other outcomes related to social and economic well-being for participants in Cohort 1. This evaluation uses a matched comparison group design which compares Cohort 1 (the treatment group) with participants from Cohort 3 (the comparison group) who had recently started the program, about 24 months after the disbursement of grants for Cohort 1. The analysis relies on a household survey conducted with both cohorts ("the impact survey").
3. **A process evaluation** to summarize the programmatic context, explore participants' experiences with and perceptions of the program, and identify facilitators and barriers to achieving the outcomes outlined in the program logic. The process evaluation draws primarily on the analysis of qualitative data from program participants in the first two cohorts and from program implementers. We complement this with a descriptive analysis of quantitative program monitoring data.

We also conducted a **cost-effectiveness analysis** to assess the value-for-money of the program and compare it to the cost-effectiveness of similar programs in low- and middle-income countries.

Figure III.1. Evaluation timeline, by cohort



Implementation: Business skills training Grant disbursement

Data collection activities: IGA survey Qualitative data Impact survey

Notes: Activities spanning multiple quarters are depicted in the quarter in which they began.

Table III.1. Research questions

Research question	Evaluation component
1. What percentage of grant recipients were actively engaged in IGAs 10 months after grant disbursement?	IGA validation
2. What were the impacts of program participation on social and economic wellbeing 24 months after grant disbursement? **a. What were the impacts on household consumption?** b. What were the impacts on household savings? c. What were the impacts on household durable asset stocks? d. What were the impacts on participants' self-confidence? e. What were the impacts on women's social and economic empowerment? f. What were the impacts on other outcomes related to social and economic wellbeing (including coping strategies and food security, sense of safety and wellbeing, school enrollment and attendance, and receipt of social assistance and social protection)? g. How did these impacts vary by subgroup (for example, by gender, refugee status, and age)?	Impact evaluation
3. What were the key elements of the program that led to achieving the desired program outcomes? a. Did participants understand and find value in the training content? Which components of training were the most valuable? b. What was the role of the cash grants in triggering the successful launch of individual businesses? c. Do participants view the program's supplementary support services and activities as valuable? How have they engaged with these support services?	Process evaluation
4. What is the community and business environment in which participants live and work? a. What are the barriers and facilitators to business growth and sustainability at the individual, household, and community levels? b. How has participants' level of community integration changed because of the program? c. How has participants' awareness and use of other social protection schemes changed because of the program? How has this influenced the achievement of the desired program outcomes?	Process evaluation

Notes: IGA = Income generating activity. Payment metrics are highlighted in **bold**.

Although this report is organized in terms of the underlying program logic in order to integrate related findings from across the three components of the evaluation, we return to the research questions and summarize the answers to them in the conclusion (Section VII).

B. IGA validation study

The IGA validation study is a quantitative descriptive study with the primary goal to assess the DIB's business metric for all three program cohorts combined.

1. Outcome definitions

The DIB agreement defined an IGA as (1) an active business, which is one that "has conducted at least one business transaction in connection with the grant received from NEF in a short reference period before

the data audit," or formal employment. The evaluation team, in collaboration with the DIB partners, further defined these terms during the evaluation framework stage (Borkum et al. 2022; **Box III.1**).

Specifically, the business metric considered the survey respondent to have an active IGA if they reported that their business was active and met at least one of the following criteria:

1. The data collection team observed a sales or purchase transaction during the survey; or
2. The respondent provided transaction records for a sale to a customer (sales transaction) or purchase of inputs (purchase transaction) that was dated within 60 days prior to the survey; or
3. The respondent reported a sales transaction within the last 60 days, and provided key details on the transaction, including the service provided or item sold, the amount and payment method, and the demographic characteristics of and their preexisting relationship with the customer;[4] or
4. The respondent reported a purchase transaction within the last 60 days, and provided key details on the transaction, including the amount of the purchase and mode of payment, the good or service purchased, and their preexisting relationship with the good or service provider; or
5. The respondent provides evidence of engaging in production-related activities in the last 60 days. This could be accomplished by showing the enumerator a sample of a product produced within 60 days, describing an upcoming event where goods will be sold, or reporting a down payment and the total expected payment for an ongoing order; or
6. Respondents did not have an active business but were formally employed as of the survey date. Formal employment required the respondent to meet the definition in **Box III.1** (self-reported) and describe details about their job, such as the job location, mode of commuting, and job schedule.

The IGA survey also captured several types of additional information that enabled us to better describe and understand the status of participants' IGAs. First, it captured information about business financial metrics, including monthly business revenues and costs (which we used to calculate a rough proxy for

Box III.1. Definition of key terms for IGA validation

Business transaction: Any of the following:

- **A sales transaction:** the exchange of a good or service for cash or kind; or
- **A purchase transaction:** the acquisition of goods or services related to business for cash or kind; or
- **Production-related activities** in preparation for an upcoming transaction, including production of goods, or receipt of a down payment.

Connection with the grant received: Based on grantees' self-reports of having a business connected with the grant.

Short reference period: Two months (60 days) prior to the IGA validation survey.

Formal employment: (1) Having an employment contract or contributing to the Jordanian social security system through their employer; (2) working at least 20 hours weekly on average over the previous month; and (3) earning monthly wages during the previous month that are equal to or greater than the minimum wage for their nationality and sector of employment.

[4] We did not seek to verify these self-reported details for this or other active IGA criteria; rather, we view the respondent being willing and able to provide these details as making it more plausible that they were reporting a valid activity. In practice, as we show in Section V, the vast majority of active IGAs were verified based on written documentation of a sales transaction.

monthly profits),[5] current business savings and debt, and the monthly amount taken out from the business for personal and household expenses.[6] Second, it captured self-reported information about business management practices that have been found to be positively correlated with business sales and profits in several low- and middle-income countries (McKenzie and Woodruff 2016, see Appendix B for more detail). Finally, it captured information about time respondents spent on their business and external financial investment in their business beyond the grant (questions added for Cohorts 2 and 3 only).

2. IGA survey data collection and analysis

Mathematica and Mindset conducted a separate round of the IGA survey for each cohort of grantees. Although the DIB defined the business metric as being measured 10 months after the midpoint of grant disbursement, in practice the survey timing has varied slightly across cohorts. This variation was necessary to avoid conducting the survey during or immediately after the month of Ramadan, given that business activity during Ramadan may be atypical. Specifically, the average survey follow-up period relative to grant disbursement was 10.3 months for Cohort 1, 11.2 months for Cohort 2, and 8.3 months for Cohort 3. Across all cohorts combined, the average survey follow-up period after accounting for differences in cohort size was 9.8 months, close to the 10 months period envisaged by the DIB.

For each cohort, we drew a random sample of 600 grantees, which was designed to be representative of the population of grantees in that cohort. We designated these 600 grantees as the primary sample and the remaining grantees as potential replacements to be randomly selected when members of the primary sample were unavailable. In our analysis, we applied cohort-level weights to reflect the overall population of grantees, given that our sample size was similar across cohorts whereas the number of grantees was greater in later cohorts. Thus, the IGA survey estimates are representative of the full population of grantees. Appendix B provides additional technical details on data collection and analysis for the IGA survey.

IGA survey respondents were 83 percent women, 30 percent refugees, and 23 percent youth; 5 percent had a disability (Appendix Table A1). Respondents' mean age was 36 years, 30 percent were heads of their household, and their mean household size was 4.6 people. NEF had provided them with grants averaging 589 JOD ($830, or PPP $2,191) to support their businesses. About 12 percent of respondents sought to improve an existing business through their grant while the remainder established a new business. The most common types of businesses were home food processing (26 percent), sales of clothes, shoes, and cosmetics (15 percent), home grocery businesses (13 percent), and home sewing and tailoring (12 percent).

[5] Revenues and costs were estimated from respondents' written records, where available (about three-quarters of respondents for revenues and two-thirds for costs), or else self-reported. Business savings and debt were self-reported.

[6] The survey questions about revenues, costs, and take-home business incomes referenced the calendar month immediately prior to the survey. Although these monthly metrics likely fluctuate from month to month, this provided a useful snapshot of the status of grantees' businesses at the time of the survey.

C. Impact evaluation

The impact evaluation compares the outcomes of participants in Cohort 1 about 23 months after grant disbursement with those of a comparison group of participants in Cohort 3, who had recently started the program at that point in time. Although the timing of outcome measurement is linked to grant disbursement, the impact evaluation includes both grantee and non-grantee participants from Cohort 1 (about two-thirds are grantees). The impact evaluation therefore measures the average impacts of the program on all participants, including both grantees and non-grantees; however, we also conducted additional analyses by grantee status.

1. Outcome definitions

The consumption metric, which is the key DIB metric linked to the impact evaluation, is defined as the average monetary value of annual household consumption. (**Box III.2** explains how consumption differs from other common measures of economic well-being, namely income and expenditure). Household consumption includes four categories: food, non-food goods and services, durable goods (such as appliances, vehicles, and electronics), and housing. We calculated the value of each of these four consumption categories for each household on an annual basis, as summarized in **Box III.3** and described in in more detail in Appendix C. We then added these values for each household to estimate the consumption metric.

Box III.2. Income, expenditure, and consumption

These concepts measure economic well-being in different ways:

- **Income** is the flow of monetary resources into the household regardless of how the household uses that money.
- **Consumption** is the value of the goods and services that members of a household enjoy. In addition to goods and services that the household directly purchases and consumes, it includes the value of consumed goods and services that the household already owns, receives for free, or produces itself; value derived from durable goods owned; and the value of housing even for homes that are fully owned.
- **Expenditure** is the monetary amount that a household pays to others for goods, services, or other things. Some expenditure on goods and services translates directly into consumption value, but some does not (for example, if the household purchases items for later consumption or durables for long-term use). Expenditure also includes categories that consumption does not, such as transfers to other households and debt repayments.

Box III.3. Overview of the consumption metric

Food	Non-food
The total value of food consumed by all household members, including food that was purchased, prepared at home, received as a gift, or as in-kind payment.	Household expenditures on transportation, health and hygiene products, utilities, tobacco products, clothing, maintenance of home and vehicles, health care, education, and other miscellaneous items.
Durable goods	**Housing**
The estimated value a household derives from using durable goods like appliances, vehicles, and personal electronic devices.	The rent paid or market value of rent for a household's current housing.

In addition to household consumption, the impact survey was also designed to measure a variety of other outcomes linked to the program logic. These included household income, savings, and debt, subjective sense of well-being, women's social and economic empowerment, use of harmful coping strategies to meet household needs in the face of limited food or financial resources, children's school enrollment and attendance, and receipt of assistance and social protection. We measured most of these other outcomes using preexisting, validated survey questions (including NEF's data collection instruments), all of which had been previously administered among similar populations Appendix Table C3 provides additional details on the definitions of these outcomes and **Box III.4** describes how we administered the survey questions on potentially sensitive topics. We also included in the impact survey a subset of questions from the IGA survey to measure a longer-term version of the business metric and business financial metrics for grantees two years post-grant.

2. Impact survey data collection

Mathematica and Mindset conducted the impact survey for a sample of program participants from Cohorts 1 and 3 between late May and late July 2024, about 23 months after grant disbursement in Cohort 1 and prior to grant disbursement in Cohort 3. Although the consumption metric was initially defined using a 24-month follow-up period for Cohort 1, it was necessary to conduct the survey at 23 months, before Cohort 3 started to receive grants. About 91 percent of Cohort 3 participants were surveyed after completing the core business training, but none had received grants; on average, respondents were surveyed within 25 days of completing the training.

Box III.4. Collecting sensitive data from vulnerable participants

The implementation and evaluation teams were concerned that collecting data on potentially sensitive topics like gender roles posed risks for privacy and response bias because other household members were often present to provide inputs on the consumption-related modules.

To mitigate these risks, the modules on food security, women's empowerment, self-esteem, and self-confidence were self-completed by the respondents using the enumerators' tablets and were restricted to literate respondents, who comprised the vast majority of respondents.

To select the survey sample, we conducted an initial statistical matching approach for all Cohort 1 and Cohort 3 participants, which sought to improve the similarity between the two groups in terms of socio-demographic characteristics available from program enrollment data. We then included all matched Cohort 1 participants and a random sample of matched Cohort 3 participants in the primary sample, given the much larger sample size for the latter. The rest of the matched Cohort 3 participants served as replacements in the case of nonresponse when sampled participants were unreachable, unwilling to participate, or unavailable.

3. Impact survey analysis approach

After data collection, we rematched the respondents who completed the survey using a similar set of characteristics as at the sampling stage to ensure that the analysis sample remained balanced in socio-demographic characteristics between Cohorts 1 and 3 given small differences in response rates by subgroup. We were also able to conduct a separate analysis of impacts by grantee status by rematching with grantee status as an additional matching variable (for Cohort 3, this was future grantee status, which was known by the time we conducted the analysis). Appendix C includes more information on the data collection and analysis approach for the impact evaluation.

The overall Cohort 1 impact analysis sample comprised 88 percent women, 35 percent refugees, and 24 percent youth; 8 percent had a disability (Appendix Table A2a). Mean age for this sample at the time of selection for the program was 37 years, about 30 percent of the sample were heads of their household, and their mean household size was 5 people. About two-thirds had received a grant from the program, averaging 565 JOD ($797, or PPP $2,102). In Appendix C, we show that this Cohort 1 impact analysis sample was broadly similar in characteristics to the full group of Cohort 1 program participants, supporting the generalizability of the findings to the latter. We also show thar the matching approach successfully resulted in well-balanced Cohort 1 and Cohort 3 analysis samples with similar demographic characteristics and baseline socio-economic characteristics, which supports the internal validity of the comparison group design. The conclusions regarding generalizability and internal validity are similar for the grantee-only analysis.

D. Process evaluation

The process evaluation was designed to document the programmatic context, explore participants' experiences with and perceptions of the program, and identify facilitators and barriers to achieving the envisaged outcomes. It drew primarily on qualitative data, comprising focus group discussions (FGDs) with participants from the first two cohorts and two rounds of semi-structured key informant interviews (KIIs) with program implementers. It also included a descriptive analysis of relevant program monitoring and evaluation data, which provides additional detail on participant characteristics, participation in training, grantee selection, and access to post-grant services.

In mid-2023 and mid-2024, Integrated International conducted qualitative data collection focused on Cohorts 1 and 2, respectively. For both cohorts, this occurred approximately 12 months after grants were disbursed. Mathematica provided training, support, and oversight of the process. Additional information on the characteristics of FGD participants and KII participants is available in Appendix D.

- **FGDs:** Integrated International conducted 18 FGDs with a total of 157 project participants (91 from Cohort 1 and 66 from Cohort 2). FGDs took place in each of the five project sites: we collected data from Cohort 1 participants in Amman, Zarqa, and Russeifa in 2023, and from Cohort 2 participants in Irbid and Kufrsoum in 2024. The FGDs explored participants' perspectives and experiences related to program outcomes, training content, cash grants, support services, and so on. Most FGDs were held separately by gender, age (youth versus adults), and grantee status to encourage participants to be forthcoming and better capture the experiences of different subgroups.
- **KIIs:** Integrated international and Mathematica conducted a total of 20 KIIs (14 in mid-2023 and 6 in mid-2024) with key stakeholders, including representatives from NEF and the CBOs involved in the implementation of the project. KIIs with field staff focused on the same sites selected for FGDs in each year, enabling us to triangulate the information in a common context. These KIIs gathered perspectives from those closely associated with program design, execution, and management, thus providing valuable context and insight into the program's operational dynamics.

Qualitative data analysis followed a systematic and iterative process designed to extract information from transcripts based on an initial set of themes related to the research questions, identify additional emerging themes, and triangulate information under each theme to draw out the key findings.

E. Cost-effectiveness analysis

The cost-effectiveness analysis was designed to provide insights into whether the DIB program was a worthwhile economic investment, both in isolation and relative to alternative programs, and to help inform further scale up and replication. The analysis involved the following steps:

- **Cost analysis:** We collected detailed data from NEF on program costs using the ingredients method, which entails categorizing expenses by cost type and specifying the years in which these costs were incurred. We also gathered information on the roles and responsibilities of various NEF and CBO staff, since this has implications for the costs of scale-up and replication in other contexts.
- **Estimating benefits:** To assess the benefits of the DIB program, we estimated two different benefit streams based on the findings of the evaluation: (1) business profits, measured through the IGA validation study and impact evaluation, and (2) household consumption, measured through the impact evaluation. We calculated these benefits for all cohorts and projected them into the future by integrating measured values from the study with estimates of other key parameters from the broader literature on micro-entrepreneurship programs in low- and middle-income countries.
- **Comparing costs and benefits:** We implemented standard approaches to account for inflation and the time value of money so that benefits and costs are directly comparable, and calculated the benefit-cost ratio separately for business profits and household consumption. Finally, we assessed whether the program's estimated cost-effectiveness is robust to different parameter assumptions about the future, and compared findings to cost-effectiveness measures from other, similar programs.

Additional details on the approach to the cost-effectiveness analysis and its findings are provided in Appendix G, with the findings summarized in Section VII.

IV. Findings on implementation and outputs

This section presents findings related to the implementation of key program activities and the related outputs in the logic model, based primarily on the process evaluation. These findings cover recruitment and selection of participants into the program, training activities and business plan development, selection of participants for cash grants and utilization of the grants, and support provided to grantees in the months following grant disbursement. They are directly related to the research question regarding which elements of the program led to desired program outcomes (Research Question 3), which requires understanding how the various program activities were implemented and how participants perceived them. **Box IV.1** summarizes the key findings.

Box IV.1. Key findings: Implementation and outputs

- NEF's **adaptive management approach led to continuous program improvement** over time. In response to feedback from Cohort 1, the program made trainings more practical, improved flexibility and accessibility to accommodate participants' other life responsibilities, provided additional support to grantees in spending their grant, and improved communication during the grant selection process.
- **The core business skills training was perceived as inclusive and highly valuable** by both grantees and non-grantees. Feedback on the life skills training was more limited, but some found value in the modules on communication, self-confidence, and public speaking. Perceptions of the technical/vocational trainings were generally less positive, with several participants expressing a desire for trainings that were better suited to their businesses and existing skills, but there was some improvement in response to feedback from Cohort 1.
- The program used a **rigorous grant selection process** to identify proposed businesses with strong potential for success. However, the **interview that was part of the grant selection process was anxiety-inducing and uncomfortable for many participants**, especially women. There is still room for future improvement to facilitate applicants' success in the grant selection process—for example, by offering more details about the selection criteria and/or by providing additional interview preparation.
- The **program's cash grants provided a strong foundation for grantees to implement their businesses plans.** It was common for grantees to invest the entire grant in their businesses, and many supplemented it with personal resources. However, the grant ceiling posed a constraint to start-up and growth for some grantees whose businesses were capital-intensive, based outside the home, and/or operating in Amman.
- **During the post-grant period, grantees benefited substantially from one-to-one mentorship**, during which trainers conducted site visits to grantees to provide refresher trainings and offer support and encouragement. ◢

A. Perceptions of training content and implementation

1. Business skills trainings

FGD respondents found the business skills trainings very useful, especially the modules on budgeting, calculations, and marketing/advertising. Other modules highlighted by FGDs respondents as being useful included those in customer service, market analysis/purchasing, identifying strengths and weaknesses, and managing risk. Most FGD respondents also found the training to be appropriately paced and logically sequenced. A few respondents struggled to keep up with the pace, but they typically attributed their difficulties to their limited literacy or education rather than issues with the training delivery, and most reported receiving adequate one-to-one support from trainers to help them keep up. Despite overall positive feedback on the business skills trainings, respondents in Cohort 1 frequently emphasized a need for more opportunities to apply the training knowledge in practice, especially in areas like marketing and purchasing. For example, many Cohort 1 respondents felt that the social media marketing training provided a strong base of conceptual knowledge but insufficient opportunities for hands-on experience using social media tools to advertise and reach new customers.

> "We still refer to the training materials on paper-based recordkeeping of revenue and costs, even today."
>
> Male grantee, Cohort 2, Kufrsoum

> "It was too little. There were only five days of training. Most of it was only theory to write the business plan. Also, when it came to marketing training, we did not get the chance to implement what we learned."
>
> Female grantee, Cohort 1, Russeifa

Cohort 2 participants found the training content to be more practical than Cohort 1 did, reflecting NEF's adaptive management and responsiveness to participants' feedback. For Cohort 2, NEF partnered with a private vendor to train trainers, improving trainers' capacity to deliver high-quality trainings that were aligned with participants' knowledge and needs. NEF also used feedback from Cohort 1 to update the training materials to make them easier for participants to understand, and to provide guidance to trainers to focus on topics that Cohort 1 participants found most challenging. For Cohort 2, trainers placed more emphasis on practical examples, supplementing conceptual explanations with videos, visual aids, and a variety of real-life examples. As a result, Cohort 2 FGD respondents found the training modules more practical and immediately applicable to their projects[7] than Cohort 1, and many participants reported referencing their training materials in their day-to-day business operations. For future programming, several FGD respondents from both cohorts suggested providing photography trainings to supplement the training on social media marketing. These respondents said that photography skills are necessary to effectively advertise products on social media and attract customers.

Based on the experience with Cohort 1, the program also implemented strategies to make trainings more flexible and accessible, helping participants balance their training schedules with other life responsibilities. In the first round of qualitative data collection, several FGD respondents and interviewees suggested providing trainings in a more flexible format to reduce absenteeism and accommodate participants with challenging schedules, like university students or people who were employed. Because several staff members had noted that female Cohort 1 participants faced particular

[7] We follow the program's terminology that referred to participants' planned businesses as "projects".

challenges balancing training attendance with childcare and other responsibilities at home, NEF established four dedicated children's areas, offered through local CBOs, to allow mothers with young children to participate fully in trainings. Staff members also worked with participants to schedule post-grant visits around participants' work and home responsibilities, or during times when they would be least burdened by childcare responsibilities, such as while children were at school. To further accommodate participants' schedules, the program provided the two-day life skills training to both cohorts through a platform called Siraj Digital, which was also used to share recordings of business trainings to support review by participants outside of scheduled sessions. The life skills training on this platform was self-paced and could be completed at any time. Finally, because the overlap of some Cohort 1 trainings with Ramadan negatively affected engagement and attendance for some participants, the program implementers avoided conducting Cohort 2 training activities during Ramadan.

During business skills trainings, implementers grouped participants with complementary projects together to encourage cooperation and knowledge-sharing. Participants undertook the business skills trainings in small groups. Program staff selected the training groups to maximize opportunities for future cross-business collaboration; the groups were diverse in terms of gender, nationality, age, ability, education level, and occupational background. A few grantees in the FGDs mentioned that they eventually partnered with other participants on business activities, indicating some degree of success in promoting cross-business collaboration. Although many FGD respondents agreed that this approach was helpful in facilitating the sharing of ideas, several respondents in both cohorts suggested grouping training participants into even more targeted subgroups (for example, by age group or project type) to enable trainers to tailor content and address the unique learning needs of each group. However, there was no consensus on this among FGD respondents, and program staff members generally agreed that trainings should be delivered to mixed classes, with trainers providing more individualized support to specific participants where necessary.

"We usually merge projects we feel would be well-suited together. For example, a participant who makes candles, another who sells flowers, another who runs events—these go together."

Trainer

"After coming to the training, I started adding more products. Through the training, I was able to collaborate with another participant to sell kibbeh [a local meat-based dish] in addition to pastries. We got new ideas on how to enhance our business."

Female grantee, Cohort 1, Amman

"It's better not to separate the classrooms by age. Youth can be energetic and active, while mature participants are calmer and more focused. Merging them in a group will maintain balance—so the room is lively, and participants are not bored nor distracted."

CBO staff

Project staff generally observed positive integration between Jordanians and Syrians, as well as between women and men, during business skills trainings. Staff emphasized that having diverse groups of participants in the training sessions did not cause significant problems and that most groups naturally got along without significant challenges. A few staff members noted that initial friction among trainees typically resolved itself, a success they attributed to the positive inter-group interaction and communication facilitated by the training sessions. Multiple staff members noted that training promoted inclusivity and allowed everyone to contribute, regardless of their nationality or gender.

> "We had participants from different nationalities as well as a mix of men and women of different age groups. There were no significant challenges in terms of integration during training. It was normal for all of us to interact and work together. Everyone treated each other with respect and understanding, and any minor issues were quickly resolved. We set common ground from the beginning, and this was beneficial for everyone to understand and respect each other."
>
> Master trainer

The project also created opportunities for supportive interactions between men and women who were in the same training groups. Although men were typically a minority in the training groups, most men who participated in focus groups said this did not make them uncomfortable and that they supported women in their training groups. Most women did not provide specific feedback during focus groups on any experiences with men in their training sessions.

Participants who were illiterate or who struggled to read and/or write received individualized support during training, which facilitated their success. According to NEF administrative data, more than 97 percent of project participants across all three cohorts were literate. However, some of those who are literate might have low levels of literacy; nearly 8 percent of participants had a primary school education or less. Staff generally emphasized that selection into the program was non-discriminatory, including in terms of literacy levels, but several staff members and participants noted that participants with more education generally found it easier to understand the training content.

A Business Development Officer we interviewed explained that NEF worked early to identify low literacy participants through the vulnerability assessment. A volunteer was then assigned to provide these participants with one-to-one support during the training. Some trainers also provided further support—for example, recording sessions through the Siraj Digital platform so that participants who struggled to keep up with trainings due to illiteracy could listen back to them later. For FGD respondents who struggled with literacy, the one-to-one support provided was critically helpful, especially when writing their business plans.

2. Life skills trainings

Participants provided more limited feedback on the life skills training, which did not appear to be central to their program experience. The life skills training, which was offered in a two-day, self-paced virtual format, included modules on gender-based violence awareness and response strategies, self-esteem, and development of soft skills such as communication, problem-solving, and time management. For Cohort 1, this training was only offered to grantees; 98 percent of Cohort 1 grantees, but only 66 percent of participants overall, completed the training, and FGD respondents did not provide feedback about it. For Cohort 2, the life skills training was offered to all participants and completed by 93 percent of

grantees and 89 percent of participants overall. Female grantees from Cohort 2 felt that the problem solving, self-confidence and communication/public speaking topics covered during the life skills training were useful. Staff members additionally suggested that the training module on gender-based violence reduction was important and impactful, but this was not mentioned by any program participants during the focus groups.

3. Technical/vocational trainings

For Cohort 1, most FGD respondents did not participate in the project's technical/vocational training, which many were unaware of, and which were not relevant to most occupations. According to NEF leadership, unexpected delays in the procurement process for technical/vocational training resulted in a shortened window for implementing these trainings for Cohort 1, which led to lower-than-expected participation rates. Cohort 1 FGD respondents reported that technical/vocational trainings were only offered for cooking and sewing, which were not relevant to many of their projects. Respondents suggested offering trainings relevant to more occupations. In addition, most Cohort 1 FGD respondents reported that they were not invited to these trainings, and many were not even aware of which trainings were being offered. They suggested advertising the technical/vocational training offerings more effectively so that participants are better aware of which trainings are available.

Cohort 1 FGD respondents who did attend the cooking and sewing trainings thought the content was not advanced enough. These respondents emphasized that most members of their cohort interested in establishing businesses in these fields had many years of relevant experience and did not learn anything new from these trainings, which focused on building foundational competencies and basic skills. They emphasized the need to offer more, and more advanced, technical/vocational trainings. For example, one participant suggested teaching cooking participants about the nutritional aspects of food-making, or how to make new types of food from other parts of the world.

NEF updated the vocational training content to be more advanced and cater to more business types for Cohort 2; however, future programs could further tailor the offerings to participants' learning goals. Responding to feedback from Cohort 1, NEF more closely examined the curricula proposed by each vendor that applied to provide the technical/vocational trainings for Cohort 2 and considered the curricula as a selection factor when selecting vendors. Cohort 2 FGD respondents generally had more positive feedback on the usefulness of the technical/vocational training than those in Cohort 1, and overall participation rates were also higher among Cohort 2. Still, the participation rate was relatively low overall for this cohort (35 percent), and some Cohort 2 FGD respondents continued to express a desire for more specialized trainings on more advanced skills, echoing the experience of Cohort 1.

B. Perceptions of grant selection, and use of grants

1. Selection of grantees

Many female FGD respondents from both cohorts said that the in-person interview required as part of the grant application process was anxiety-inducing and uncomfortable. The program used a rigorous grant application process that comprised two components: (1) the evaluation of the written business plan developed during the business skills training and (2) an interview. The business plan was evaluated by a Business Development Officer and contributed 60 percent of selection. The interview was worth 40 percent of selection. It was conducted by members of a selection committee, who asked questions about the business plan and could also ask for supplementary information, such as pictures of the products an applicant hoped to sell. Through this rigorous process, the program sought to focus grants on business plans that had strong potential to succeed.

> "The interview was quite tough. I hadn't prepared photos of my work since my project is still new. The interviewer insisted on seeing them, which made me feel uneasy."
>
> Female grantee, Cohort 1, Russeifa

Several non-grantees who participated in the FGDs said that their nervousness made them unable to present their business plan effectively or answer the interviewers' questions. These respondents would have liked the program to create a more relaxed environment for interviews. One group of program staff further suggested training the interview committee on listening and evaluation skills that would put participants more at ease. Although not mentioned by FGD or KII respondents, including a short training module specifically to prepare applicants for the interview (for example, covering some basic interview-relevant soft skills) might also be helpful. Some staff members and non-grantees also suggested supplementing the interviews with field visits to applicants' businesses (if they already existed or were in the process of launching) to allow applicants who struggled to express themselves during the interview to demonstrate their strengths and improve the committee's understanding of their projects.

FGD and KII participants proposed several ideas to improve applicants' chances of success in the grant application process, in addition to reducing the stress of the interview itself. An NEF Business Development Officer suggested that trainers show examples of successful projects and explain which other types of projects might face obstacles, such as legal/regulatory requirements and high start-up costs. One participant suggested adding a preliminary review stage to the grant application/approval process, where the committee would screen applicants, identify those whose ideas had a lot of potential but needed further development, and refer them for more ideation support from trainers. Staff members interviewed in the second round suggested making additional improvements to communication to improve the transparency of the selection process**,** such as providing even more details about the selection criteria. For example, this could mean ensuring that applicants are aware of the level of market saturation and competition in their project area, which may make it less likely that their business will to succeed, and in turn reduce their chances at being selected for a grant.

Communication to rejected grant applicants improved based on feedback from Cohort 1. Many non-grantees in Cohort 1 who participated in the FGDs were frustrated by poor communication and a lack of feedback on their business plans throughout the grant selection process. According to CBO staff and

Siraj Officers, all rejected applicants from Cohort 1 were supposed to receive an automated WhatsApp message explaining the reason for their rejection within a month of the decision. However, none of the Cohort 1 non-grantees who participated in the FGDs recalled receiving any messages. Because many non-grantees continued or plan to pursue their businesses without grants, they would have greatly valued detailed feedback on their applications even if not awarded grants. Other communications during the application process for Cohort 1 felt vague to some non-grantees and made them feel confused about the status of their candidacy.

Based on this feedback, NEF implemented improvements for Cohort 2. NEF staff members explained that the program began to communicate the reasons for rejection, and more (but not all) Cohort 2 non-grantees who participated in the FGDs confirmed that they received a reason for their rejection. Rejected applicants were also encouraged to re-apply for a subsequent cohort, which at least one Cohort 2 non-grantee who participated in the FGDs said inspired them to do so.

There was a perceived point to a lack of clarity among some Cohort 1 participants about the evaluation criteria for grants, but this improved for Cohort 2. Implementers emphasized that the grant application process was fair and non-discriminatory. However, Siraj Officers noted that the project was legally unable to approve businesses which had to go through licensing/legal procedures to open a shop (the program focused primarily on home-based businesses or businesses that already existed as shops). Some Cohort 1 non-grantees who participated in the FGDs believed that the committee gave preferential treatment to certain applicants based on their project location, project type, or gender. For Cohort 2, improved communication and transparency may have supported an improved perception that the application process was fair. Cohort 2 non-grantees who participated in the FGDs had a clearer understanding of the evaluation criteria, and more of them received detailed explanations for their rejections. Although a few Cohort 2 non-grantees who participated in the FGDs still questioned whether the true reason for their rejection was bias against them based on their gender, age, or literacy level, they were generally less likely to perceive the selection process as unfair.

> "It's essential for master trainers to inform participants about successful projects and those that may face obstacles, like projects with high financial requirements or those not feasible within our context due to regulatory barriers. Being transparent about potential challenges can help participants make informed decisions about their projects."
>
> NEF Business Development Officer

Quantitative data suggest that the merit-based grant selection process may have inadvertently favored some demographic groups and led to the selection of grantees who were somewhat more advantaged. An analysis of NEF's vulnerability assessment data showed that representative proportions of men and women were selected for grants, but refugees or youth were less likely to be selected than Jordanians or adults (by 5 and 13 percentage points, respectively) (**Table IV.1**). Further, grantees had baseline asset levels that were 0.11 standard deviations higher than non-grantees, a modest but statistically significant difference. We also used the impact survey data, which were collected after Cohort 3 completed training but before they received grants, to compare additional socioeconomic characteristics of Cohort 3 grantees and non-grantees. Cohort 3 grantees reported average household incomes that were 14 percent higher than non-grantees, although consumption levels were similar. Taken together, these findings suggest that some groups were better positioned to succeed in the business plan development and grant selection process, even if the process itself was merit-based.

Table IV.1. Characteristics of grantees and non-grantees

Characteristic	Grantees	Non-grantees	Difference
Vulnerability assessment (all cohorts, at program entry)			
Women (%)	80.9	81.2	-0.3
Refugee (%)	30.0	34.9	-4.8**
Youth (%)	25.7	38.9	-13.2**
Asset index (standard deviations)	0.05	-0.07	0.11**
Sample size	**3,416**	**2,244**	**n.a.**
Impact survey (Cohort 3 only, before grants awarded)			
Household income	4,200	3,687	513*
Household consumption (JOD)	6,455	6,317	138
Sample size	**451**	**266**	**n.a.**

Source: NEF Vulnerability assessment and the impact survey

Notes: n.a. = not applicable

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test;

2. Use of grants

Grantees primarily used the grant to purchase the equipment, appliances, tools, and/or raw materials they needed to start new businesses or expand existing ones, per their business plans. FGD respondents generally reported investing their entire grants in business expenses, with some grantees from Cohort 2 additionally reporting setting aside a small amount as savings for emergencies. Most Cohort 3 IGA survey respondents reported spending their grants on business-related equipment (85 percent of grantees), raw materials or other inputs for products they make or services they provide (for example, ingredients for home food processing, or hair products for salon services; 63 percent), and purchasing preexisting or wholesale inventory for resale (17 percent). (We did not collect this information in the IGA surveys for Cohorts 1 and 2.) NEF leadership suggested that grantees were encouraged to set aside up to one-quarter of the grant for other necessary expenses beyond initial purchases, such as transportation and electricity; however, some grantees struggled financially because the amount

"This grant was a game-changer for my future plans. It helped me prioritize the equipment and tools I absolutely need to get started, and then I could focus on researching the best quality materials for my business."

Male grantee, Cohort 2, Kufrsoum

"Did I pay anything towards my house? No, on the contrary—I spent more of my *own* money on the project."

Female grantee, Cohort 1, Amman

they set aside was insufficient to cover expenses that were either unanticipated (such as replacement costs for faulty equipment) or greater than expected (including electricity costs and rising operating costs due to inflation).

To invest in their businesses, many participants also combined the grant funding with investments from their personal savings, as well as other forms of capital they already owned, such as equipment or vehicles. A few participants borrowed additional money or took out loans alongside the grant to invest in their businesses. For other participants, the grants provided a substitute for less sustainable or desirable financing sources, such as loans from family members or neighbors. In the Cohort 2 and 3 IGA surveys, which covered a representative sample of grantees, 50 percent of respondents reported investing external funding in their business; among these respondents, the average amount of external investment was 369 JOD ($520, or PPP $1,373). The most common expenditure categories for additional funds among Cohort 3 grantees were similar to the most common expenditure categories for the grants, although with a greater emphasis on business-related equipment rather than raw materials. (We did not collect information on external funds for Cohort 1, or on how these funds were spent for Cohort 2.)

In making initial business purchases, grantees faced some challenges in finding quality products at fair prices. A common challenge faced by grantees in both cohorts was that some shops refused to provide written price quotations or receipts. Respondents explained that some vendors do this to avoid paying taxes on the sale. Without receipts, grantees struggle to submit proof of their purchases in a timely manner, which the project required. A lack of documentation of purchase caused additional challenges for a few grantees—for example, one grantee encountered trouble with law enforcement when trying to transport their purchased items to different governorates without proof of payment. Grantees from Cohort 2 further explained that some stores exploited project participants by charging fees for quotations during the business plan phase of the training. Although NEF leadership suggested that grantees typically had around one month to make purchases and submit receipts, Cohort 1 grantees from Amman reported that the project only gave them one week to purchase materials, which was not enough time to gather information, compare vendors, and make wise decisions. As a result, several reported purchasing low-quality or overpriced products that inhibited their operations and growth. NEF staff explained that the program responded to the feedback from Cohort 1 in Amman and worked to ensure that grantees at all sites were allotted sufficient time for purchasing in subsequent cohorts, as intended.

"Some vendors refuse to provide receipts or a price quote because of taxes. I regularly take some of my products to Irbid to sell there, and on the way the police stop cars and search them. If I do not provide receipts, they will think I am smuggling something! But what am I meant to do if I am not given a receipt?!"

Female grantee, Cohort 1, Amman

"We were in a hurry to buy everything we need. If we had more time, I might have been able to ask around and buy better quality products. I'm also working simultaneously, which made things difficult."

Female grantee, Cohort 1, Amman

"I purchased a faulty product, and I could not return it. It was a waste of money and made everything more difficult. In the future, I'd ask around to make sure any product I purchase is of good quality. I needed more time to explore the market."

Female grantee, Cohort 1, Amman

Although all grantees agreed that the grant provided a good foundation for their business plans, the grant ceiling posed a constraint to start-up and growth for businesses which were capital-intensive, based outside the home, and/or operating in Amman. Many grantees had additional ideas for how to expand or improve their operations, which the grant did not cover. These ideas included investing in automated/electric equipment, replacing broken machines, or opening a shop outside the home to expand operations.

Male FGD respondents across all sites were more likely than female grantees to voice concerns about the grant amount. This might be because men were more likely to receive grants for capital-intensive businesses, like mechanical repair or blacksmithing. Many of these grantees said that the grant was not sufficient to cover the costs of equipment required to start a proper business in their field and/or open a physical shop.

Additionally, FGD respondents in Amman were more likely than grantees at other sites to voice concerns about the grant amount. In Amman, many participants had to cut down on the scope of their original business plans because their plan was too expensive given the high price levels there. Staff members explained that businesses in Amman face the greatest barriers to success due to the high level of competition and the large size of the community, which makes it difficult to reach customers.

> "In my field, I'm sometimes called to fix villas. For that, I need money to buy additional equipment. I don't have capital, so I can't take these jobs."
>
> Male grantee, Cohort 1, Zarqa

> "My original plan needed 1,000 JOD. I showed this to the supervisor, who told me to take some things away and make do. It worked, but I still need those things to do better."
>
> Female grantee, Cohort 1, Amman

C. Perceptions of post-grant support

Grantees largely appreciated the one-to-one mentorship, which was the main form of post-grant support provided. NEF provided post-grant support for all three cohorts, focusing on ongoing monitoring and one-to-one mentorship; starting with Cohort 2, they introduced a live dashboard to track and respond to participants' needs. One-to-one mentorship was typically conducted through site visits to grantees' home by a trainer. Grantees who participated in FGDs generally valued these post-grant visits and found them especially helpful in providing support with implementing the training content—for example, help with budgeting and business calculations—although a few found the visits repetitive. To improve post-grant visits in the future, a few grantees suggested that the individuals conducting post-grant visits should have specialized training in the grantee's project area to provide grantees with more tailored support and advanced follow-on training.

> "We continued to provide one-to-one coaching through follow-up visits every month. During these visits, we addressed any specific needs or challenges the participants had, and we even conducted additional training sessions as required and helped them with networking."
>
> Master trainer

> "The visits were used to cover missing information. We could call program staff for advice. They visited to check up and follow up with us. It motivated us to do better."
>
> Male grantee, Cohort 1, Zarqa

In addition to the post-grant visits, some participants received in-kind post-grant support. Trainers additionally explained that the Siraj Center provided equipment and advertising materials during the post-grant period. While most grantees who participated in FGDs were satisfied with the in-kind support they received, others said it was insufficient or irrelevant. For example, several Cohort 2 participants received equipment that they could not use (like a cordless drill for a project that did not involve construction, or a packaging machine for the wrong type of good), and at least one participant was unable to exchange the product for cash to use to buy the correct equipment.

Other aspects of post-grant support, including advanced grants and additional financing, small-group coaching, and peer business networking events were not explicitly addressed by FGD respondents. The advanced grants and additional financing aspects of post-grant support were not covered during FGDs because very few participants received them. FGD respondents also did not mention small group coaching, which does not appear to have been central to their post-grant experience. Although peer business networking events were ultimately not offered for Cohort 1 and only a handful of events were offered for Cohort 2, several respondents pointed to the social connections they formed with other participants during the trainings as a positive networking outcome, enabling them to collaborate and exchange information subsequently. For example, Female participants from Russeifa have an informal WhatsApp group, which they frequently use to meet up with each other and share advice.

V. Findings on short-term outcomes

In the short term, the program logic anticipated that participants would build their knowledge and technical skills through trainings and grow in their self-confidence. Then, by leveraging their newly developed skills, cash grants, and ongoing program supports to overcome early obstacles, grantees' business would survive and grow. Alternatively, they could use their increased skills and confidence to succeed in the workforce. To assess whether these changes were realized, this section integrates findings from the process evaluation and IGA validation study on the short-term outcomes experienced by participants within the first year after grant disbursement. We begin by examining effects on the skills, knowledge, and self-confidence of participants that were intended to provide a foundation for successful IGAs. We then examine effects on short-term business survival and income (Research Question 1), as well as the facilitators and barriers related to these outcomes (part of Research Question 4). **Box V.1** summarizes the key findings.

Box V.1. Key findings: Short-term outcomes

- The **business skills training helped participants develop critical skills** to successfully establish and operate their businesses. Grantees put these skills into practice, and most reported **implementing small business management best practices** that are typically associated with other positive business outcomes.
- FGD respondents reported a **greater sense of self-confidence, motivation, and independence** as a result of the training.
- **About 10 months after grants were disbursed, almost all grantees' businesses were still active;** the vast majority were conducting frequent transactions and earning positive profits. Mean monthly profits were 133 JOD ($188, or PPP $495), of which a mean of 98 JOD ($138, or PPP $365) was take-home income that went towards supporting personal and household expenses. The latter is equivalent to about one-third of the national minimum wage and about one-third of mean monthly household expenditures for refugees.
- **Male grantees reported higher levels of revenue, profits, and take-home income from businesses than female grantees.** This is likely related to differences in business types, the additional resources they have invested in their businesses, and the amount of time they spend each week on their businesses.
- **Key facilitators to business survival and growth were family, community, and program-provided post-grant support, while key barriers included rising costs, competition, and challenges with business registration.** Refugees and female grantees face additional barriers and constraints to operating and growing their businesses.

A. Increased skills, knowledge, and self-confidence

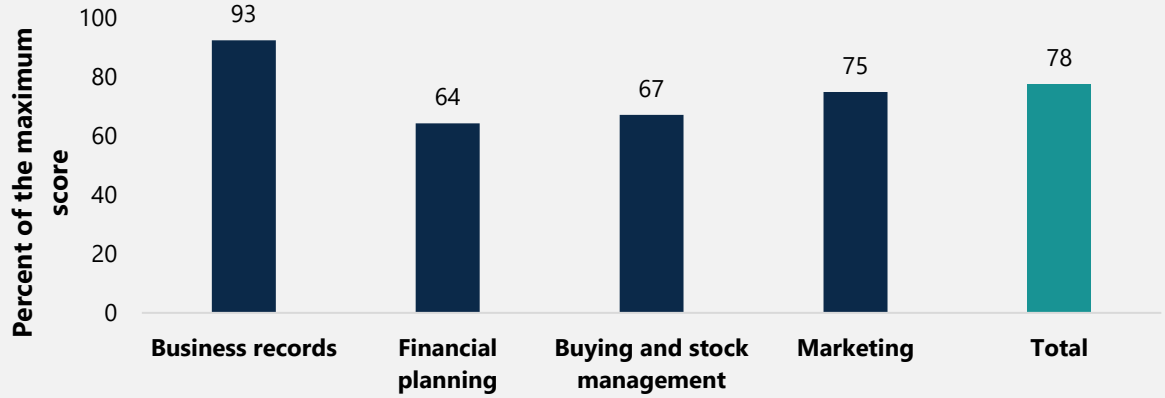
1. Business skills and practices

Grantees reported implementing business management skills that are associated with positive business outcomes. Both grantees and non-grantees who participated in the FGDs felt that the training had helped them build critical business skills. One year after training, grantees reported that they were implementing many of these skills to reduce costs and better keep track of their expenses. Even many non-grantees who participated in the FGDs reported that they were still pursuing their business plans—continuing to operate a pre-existing business or planning to start a new small business in the future—and they felt that the trainings had prepared them to pursue those activities.

> "I didn't used to keep registers or books. Now I keep track of my businesses' money. It has made a massive difference, knowing where everything is going."
>
> Female grantee, Cohort 1, Russeifa

Drawing on a business practice measurement approach that was developed by researchers at the World Bank, we also used the IGA survey to assess the extent to which grantees reported implementing business management best practices in four different domains: (1) business records, (2) financial planning, (3 buying and stock management, and (4) marketing. These scores have been found to be positively correlated with business sales and profits in several other low- and middle-income countries (McKenzie and Woodruff, 2016; see Appendix B for details). Overall, grantees who responded to the survey scored 78 percent of the maximum possible score, on average (**Figure V.1**). They reported especially strong practices related to business record-keeping (93 percent, on average), which the program emphasized. They also performed well on the marketing index (75 percent), with slightly low scores for financial planning practices (64 percent) and buying and stock management practices (67 percent). Business practice scores were largely similar across the three cohorts (Appendix Table A7). There were few substantive differences in business practices between subgroups (Appendix Table A4).

Figure V.1. Average business management practice scores for all cohorts (percent)



Source: IGA surveys

Notes: The sample is weighted to account for differences in cohort size. Scales are adapted from McKenzie and Woodruff (2021). The scales have different numbers of items: 7 for business records, 4 for financial planning, 3 for buying and stock management, and 7 for marketing, for a total of 21. The total score reflects the sum of the raw scores across all scales divided by the maximum possible score of 21.

Evidence for other improvements in skills and knowledge—including technical/vocational skills and knowledge of business legal requirements—was mixed and were not central to the program's causal chain. Overall, only a minority of grantees participated in the technical/vocational component of the trainings, and few who did found these trainings to be transformative. Although highlighted in the program logic, improving grantees' knowledge of business legal requirements was also not central to the program's results because the program primarily focused on home-based businesses, which are generally exempt from business license requirements in Jordan. Moreover, participants operating businesses based outside the home reported facing license-related difficulties related to factors outside of the scope of the program (for example, laws barring refugees from obtaining a business license), rather than to grantees' lack of knowledge.

2. Access to markets and suppliers

The program helped participants navigate markets and relationships with suppliers and customers more successfully. The program logic suggested that access to markets and suppliers by program participants was important to support business survival and longer-term business success, FGD respondents indicated that the training program helped improve their knowledge and skills related to assessing offers and price quotations from suppliers, as well as building relationships with customers. Although we do not have data on the number of suppliers or access to markets, the IGA survey provides evidence on related practices. Across all cohorts, IGA survey respondents reported close engagement with their suppliers. About three-quarters of respondents reported that they had attempted to negotiate prices with a supplier and a similar fraction had asked a supplier about promising or in-demand products with in the last 3 months. Respondents were also attempting to increase their access to customers and markets; about 9 in 10 reported conducting some kind of advertising within the last 6 months, and a similar fraction reported using a special offer to attract customers within the last 3 months.

3. Access to support services

Only a few FGD respondents reported becoming involved with other CBO services following participation in the program. The program logic posited that by connecting program participants to Siraj centers, the program might increase participants' awareness and use of other social protection services. Respondents in a few of the FGDs reported learning of other social assistance programs through their association with the Siraj Center, including food distribution programs, other entrepreneurship and training programs, medical care supply distributions, pop-up health clinics, and programs to support female schoolteachers. However, respondents did not report increasing their utilization of such services after participating in the program—in fact, respondents in one FGD emphasized that participating in the small business project made them less likely to utilize these services by increasing their self-reliance and financial independence. However, a few respondents reported becoming more involved in other CBO services after the program—in particular, by pursuing additional training courses.

4. Short-term effects on self-confidence, motivation, and independence

The grants and associated business income boosted grantees' sense of self-confidence, motivation, and independence. Several FGD respondents indicated that their sense of self-confidence increased as a result of participating in the program. The training sessions made these respondents more confident in their business, communication, and problem-solving skills, and the positive feedback and continuous mentorship they received from trainers helped to further strengthen their self-confidence during the post-grant period. Grantees commonly pointed to increased financial independence offered by micro-entrepreneurship activities as a key benefit of the program, and many respondents tied increased independence to a greater sense of self-confidence.

> "Participating in the small business project activities made me more confident in myself. I gained greater knowledge about how to develop my project and exploit the opportunities available in my community."
>
> Male grantee, Cohort 2, Kufrsoum

During FGDs, grantees generally displayed a positive outlook towards the future. Several grantees who reported that participating in the program increased their motivation to build a better life for their families, persevere through adversity, and pursue their business goals. Most non-grantees who participated in the FGDs said that they felt motivated to participate in further training and continue to pursue their business plans.

B. Short-term business survival and income

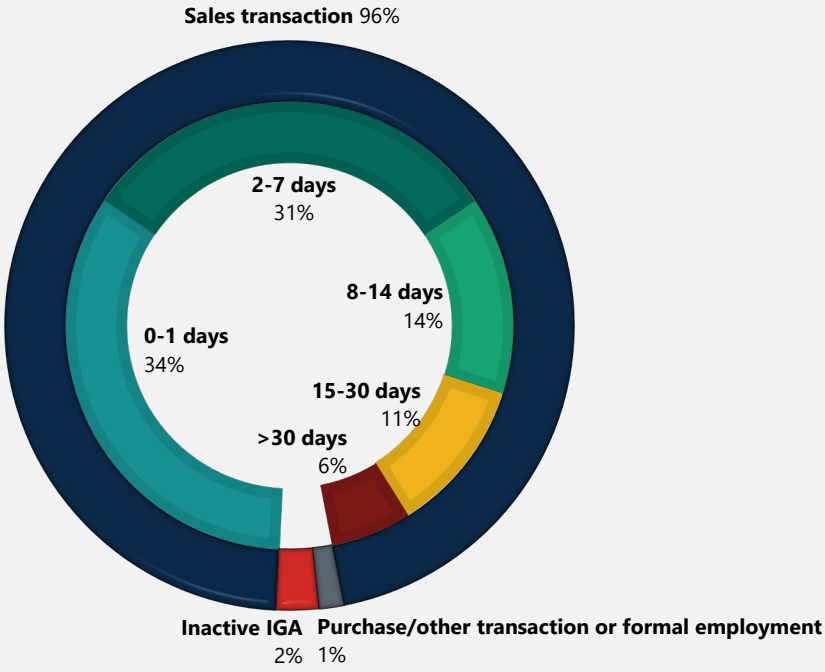
1. Short-term business survival

About 10 months after grant disbursement, 98 percent of grantees had an active IGA. This business metric was similarly high across both cohorts (Appendix Table A7) and across demographic subgroups (Appendix Table A5).[8] It substantially exceeded the threshold at which the maximum payments to investors were made under the DIB (75 percent), which was set based on findings from previous iterations of the program. In addition to learning from the previous program iterations, NEF leadership has noted that the DIB funding model has encouraged them to innovate, implement adaptive performance management, and invest in internal systems that may have enhanced overall performance.

In the short-term, grantees were conducting frequent sales transactions and tracking them carefully. Almost all respondents who satisfied the criteria of the 10-month business metric did so by verifying that they had conducted a sales transaction within the past 60 days (**Figure V.2**). To verify a sales transaction, the vast majority provided written documentation of a recent transaction. About two-thirds of respondents had conducted at least one such transaction within the 7 days prior to the survey, including 2 percent who conducted one during the survey. Only 0.2 percent of respondents satisfied the business metric based on the employment-based criteria; the vast majority of the percentage of respondents who met the criteria for formal employment also had active businesses.

[8] The 95-percent confidence interval around this estimate is 97.2 to 97.9 percent, meaning that there is a 95 percent probability that the business metric for the population of grantees is between 97.2 percent and 97.9 percent.

Figure V.2. Business metric for all cohorts, by type of evidence provided and timing of the most recent sales transaction



Source: IGA surveys

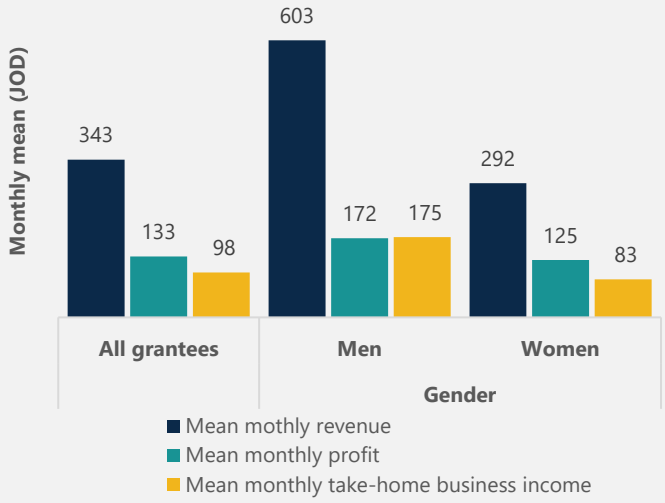
Notes: The sample is weighted to account for differences in cohort size. The business metric of active IGAs (98 percent) comprises grantees with a recent sales transaction (96 percent) plus those with a recent purchase, other business activities, or formal employment (1 percent). Totals may differ due to rounding.

2. Business profitability

Most respondents with active businesses reported that their businesses were profitable during the calendar month preceding the survey. Mean revenues for active businesses in the full month prior to the survey were 343 JOD ($484, or PPP $1,276) and mean costs were 210 JOD ($296, or PPP $781) (**Figure V.3**, Appendix Tables A6a–A6b). Nearly 9 in 10 grantees reported revenues that were higher than their monthly costs, translating into mean estimated monthly profits of 133 JOD ($188, or PPP $495) and median estimated monthly profits of 100 JOD ($141, or PPP $372). Grantees reported modest overall business savings, with a mean of 226 JOD ($319, or PPP $841) and a median of 100 JOD ($141, or PPP $372). They also reported relatively low overall business debt, with a mean of 75 JOD ($106, or PPP $279) and a median of 0 JOD.

Male grantees reported higher levels of revenues and profits than female grantees. Male grantees reported mean revenues of 603 JOD ($850, or PPP $2,243), which translated into 172 JOD ($243, or PPP $640) in mean monthly profits after accounting for costs (**Figure V.3**, Appendix Tables A6a–A6b). Female grantees reported mean revenues of 292 JOD ($412, or PPP $1,086), less than half those of male grantees. However, their costs were also lower, resulting in more comparable but still lower mean profits of 125 JOD ($176, or PPP $465). Differences in revenues and costs between men and women might be related to differences in business types, the additional resources they have invested in their businesses, and the amount of time they spend each week on their businesses. Whereas almost half of female grantees operated home-based food processing, trade of clothing and cosmetics, or tailoring businesses, the type of businesses operated by male grantees were more variable and included business types that are likely to involve higher-value transactions (for example, mobile maintenance services). As we show below, men also tended to invest more outside financial resources in their business and spend more time on their business than women, which might be reflected in having businesses that are more active and generate more value relative to women. Differences in these business financial metrics at the 10-month mark were more modest for other subgroups (Jordanians versus refugees, youth versus adults, and across baseline asset quintiles) and they were also similar across cohorts (Appendix Table A7).

Figure V.3. Business financial metrics for all cohorts, overall and by gender



Source: IGA surveys

Notes: The sample is weighted to account for differences in cohort size.

3. Take-home business Income

The IGA survey also asked respondents how much business income they used for personal and household expenses during the full month prior to the survey.[9] This short-term outcome, measured about 10 months after grant disbursement on average, is an important measure of whether grantees' IGAs are likely to translate into medium- and long-term outcomes anticipated by the program logic, such as increased consumption. About 90 percent of respondents with active businesses reported a positive take-home business income in the month prior to the IGA survey. The mean take-home business income among all active businesses was 98 JOD ($138, or PPP $365), and the median was 55 JOD ($78, or PPP $205) (**Figure V.3**, Appendix Tables A6a–A6b). The mean amounts to about one-third of the Jordanian minimum monthly wage of 260 JOD ($367, or PPP $967). For refugees, it amounts to about one-third of mean monthly household expenditures for out-of-camp refugees, which was estimated in a 2023 UNHCR survey

[9] Take-home business income was self-reported and was not linked directly to estimated profits.

as 321 JOD ($453, or PPP $1,194).[10] Like with revenues and profits, mean take-home incomes were substantially higher for men than women—about double—but differences were more modest for Jordanians and refugees, youth and adults, across baseline asset quintiles, and across cohorts (Appendix Tables A6a–A6b and A7). Although paid employment was included as an additional potential channel for increased income in the program logic, this was limited in practice—according to the IGA survey, only 7 percent of grantees were employed and only 3 percent were employed in formal jobs that were the focus of the DIB.

4. Facilitators of business survival, growth, and income

Key facilitators identified by grantees who participated in FGDs included support from their family and community, as well as post-grant support from the project. Many FGD respondents said they enjoy substantial family and community support and only a couple said that the grant has caused neighbors or family members to treat them with envy. Beyond the follow-up visits which are formally part of the program, some trainers have made personal purchases from participants' businesses. Several grantees expressed that their trainer purchasing their products made them confident and helped them succeed. Implementers also suggested that networking and collaboration opportunities facilitated by the training helped participants to succeed by forming business partnerships with each other and sharing knowledge across different business areas, even though formal peer business networks were not created as initially planned.

> "I started encouraging other women to work and get their own money. I helped my neighbor start her own business—I am well-known in the community now!"
>
> Female grantee, Amman

Male grantees invested more resources and time in their businesses, potentially translating into stronger business performance. Grantees' investments in their business in terms of money or time might also plausibly help facilitate stronger business performance. To assess this, we collected information about outside financial investment in the business and weekly time spent by grantees on their business for Cohorts 2 and 3. For these cohorts, the average grant amount was similar between men (602 JOD; $849, or PPP $2,239) and women (595 JOD; $839, or PPP $2,213). However, 65 percent of men but only 47 percent of women reported that they had invested additional funding in the business besides the grant, largely from their own or other household members' savings (**Table V.1**). This translates into nearly four times the mean outside investment in the business, with men spending 424 JOD ($597, or PPP $1,576) and women spending 138 JOD ($194, or PPP $513), on average. Most Cohort 3 grantees reported spending the external funding on raw materials and other inputs, business-related equipment, and pre-made inventory for resale. Male and female grantees who invested external funds were similarly likely to report spending the additional resources on business equipment (88 and 84 percent, respectively), but male grantees were more likely to invest the original grant in equipment (64 percent of male grantees versus 49 percent of female grantees). This likely reflects the more capital-intensive types of businesses many men launched, which would have required the large amounts available from the grant.

In terms of time spent on their business, men in Cohorts 2 and 3 reported spending a mean of 41 hours working on their business during a typical week whereas women reported a mean of 24 hours. This

[10] UNHCR collected these data from a representative sample of refugees across Jordan in Q2 2023. https://data.unhcr.org/en/documents/details/103118.

difference is likely because women are balancing working on their businesses with childcare and other household responsibilities. Although Cohort 3 male grantees were somewhat less likely to receive help and support from their families than their female counterparts (42 versus 68 percent), both reported receiving about 10 hours' worth of assistance per week in total across their various family members.

The association between business financial performance and outside investments is weaker for other subgroups. For example, in Cohorts 2 and 3, Jordanian grantees and those from the top wealth quartile invested around 25 and 60 percent more in external funds in their businesses than refugee and bottom-quartile grantees, respectively (**Table V.1**). Yet business revenues and profits were similar across these subgroups. These findings suggest that a complex interplay of factors might explain variation in business financial performance across subgroups. Nevertheless, differences in having any outside financial investment and in time spent on the business were strong for men versus women, suggesting that those two factors might still contribute to explaining the substantial differences in business financial performance by gender.

Table V.1. Business resources available to active businesses (Cohorts 2 and 3 only)

	Sample size	Hours spent on business in a typical week	Any external funding for business (%)	Mean external funding for business (JOD)	Mean monthly revenue (JOD)	Mean monthly profits (JOD)	Mean current monthly take-home income (JOD)
By gender:							
Women	996	24	47	138	292	125	83
Men	204	41	65	424	603	172	175
Youth	282	28	55	183	334	106	87
Adult	918	27	49	187	346	141	101
Refugees	360	24	47	159	329	144	104
Jordanians	840	28	52	198	349	128	95
Quartile 1	315	26	46	150	326	137	104
Quartile 2	298	28	51	201	353	154	99
Quartile 3	286	28	53	151	370	127	102
Quartile 4	301	26	52	243	323	111	86

Source: Cohorts 2 and 3 IGA survey

Notes: We did not collect data on business resources in Cohort 1. The asset index is based on housing characteristics and durable goods ownership from NEF's vulnerability assessment, using a principal component analysis for all three cohorts.

5. Barriers to business survival, growth, and income

Challenges to future growth and business sustainability are variable across location, business type, and gender; they include rising business costs, tough competition, and challenges with business registration. FGD respondents in Amman commonly mentioned rising business costs and tough business competition as challenges to future growth. Tough competition may interact with rising costs to produce unique financial challenges for participants operating capital-intensive businesses in high-cost areas—their operating costs rise, but they cannot raise prices without losing customers to lower-priced competitors. Some grantees, especially those who wish to operate outside the home also mentioned that they are legally required to register their businesses with the government and that not doing so exposes them to some risk. However, they have not registered because the registration process is difficult, and their operations will become more expensive due to taxes. Additional support from the Siraj Centers might help them navigate the legal compliance process.

Syrians face additional financial challenges engaging in IGAs due to their refugee status. NEF staff explained that Syrian refugees are legally restricted from obtaining a business license in Jordan, restricting business growth beyond small home-based businesses. They may face differential access to finance; according to multiple FGD respondents, only Jordanians are eligible to take out formal loans, which they can invest in expanding their businesses. Refugees' access to other sources of income like wages or support from family may also be more limited and less secure. For example, one FGD respondent who relied on UNHCR payments said these payments were paused while they were applying for the grant. Another explained that their spouse has been unable to secure employment due to lack of a Jordanian work permit, leaving their small business as the family's primary income source and increasing the pressure to succeed. Finally, some grantees reported that landlords may raise rents if they learn that tenants are operating a home-based IGA. Refugee participants, who are far more likely to rent their homes than Jordanians, are particularly vulnerable to this.

Cultural norms may constrain some women-operated businesses. As mentioned earlier, female grantees tended to spend less time on their businesses than male grantees according to the IGA survey—likely because of their traditional home-making responsibilities. Differences in the types of businesses operated by women versus men—with women focusing on a limited set of business types that typically have lower-value transactions—are also likely driven in part by societal gender norms. Some FGD respondents suggested that there may be additional challenges for some women business owners in achieving high transaction volumes because of cultural norms in the community. For example, one young female grantee explained that because she is not able to let men in her house, she relies on delivering orders to her customers, which is more challenging than selling products for pick-up by customers.

VI. Findings on medium-and long-term outcomes

The program logic anticipated that the short-term outcomes discussed in Section V would lead to a sustained increase in participants' income, which would translate into reduced poverty through increased consumption and savings and reduced debt levels. As a result, households would be better able to meet their basic needs. In the longer term, these changes were expected to contribute to broader transformations both at the household level and at the community level. At the household level, increased contribution to household income for women was expected to increase their social and economic empowerment. At the community level, increased economic opportunities for both refugees and vulnerable Jordanians were expected to lead to improvements in overall community welfare in terms of living conditions, investments in education, and social cohesion.

To assess whether these changes were realized, this section focuses on the findings from the impact survey, which was conducted almost two years following the disbursement of grants to selected Cohort 1 participants. It also incorporates relevant findings from qualitative data collected from Cohort 1 about one year after grant disbursement. We begin by examining the status of Cohort 1 grantees' IGAs at the two-year mark, which provides a foundation to understand impacts on well-being (the focus of Research Question 2). We then estimate impacts on household poverty and other measures of well-being, first for Cohort 1 participants as a whole (including grantees and non-grantees) and then separately by grantee status. **Box VI.1** summarizes the key findings.

Box VI.1. Key findings: Medium- and long-term outcomes

- **About two years after grant disbursement, 76 percent of Cohort 1 grantees still had an active business.** The vast majority of these active businesses were generating positive monthly profits and mean monthly take-home business incomes were similar to those reported after 10 months.
- **Cohort 1 participants reported higher average household incomes than matched Cohort 3 participants**, primarily driven by higher rates of business ownership and associated business incomes. Mean self-reported annual household income for Cohort 1 was 17 percent (0.24 standard deviations) higher than for Cohort 3.
- **Annual household consumption was higher for Cohort 1 than matched Cohort 3 participants, mainly driven by greater food and non-food consumption.** The impact on household consumption was equivalent to 10 percent (0.22 standard deviations). There was also a modest reduction in the use of harmful food- and livelihoods-related coping strategies by Cohort 1, although use of these strategies remained common.
- **Impacts on consumption were largest for youth, Jordanians, and households with more baseline assets**, although intersectionality across subgroups and the interplay with income impacts are complex.
- **There was no quantitative evidence of impacts on self-confidence or life satisfaction, or on women's empowerment outcomes**. However, qualitative evidence suggests there have been some positive changes that the impact survey did not capture. There was also a modest impact on enrollment in secondary education.
- **A separate analysis incorporating grantee status shows that impacts were greater for grantees and minimal for non-grantees**, suggesting that program impacts were largely driven by the grants. Consumption for Cohort 1 grantees was 15 percent (0.36 standard deviations) higher than future Cohort 3 grantees. ◢

A. Medium-term business ownership and income

1. Medium-term activity of grant-funded businesses

Almost two years after grant disbursement, 76 percent of Cohort 1 grantees still had active IGAs. In the impact survey, 87 percent of Cohort 1 grantees reported that they still had a business connected to their grant. However, only 76 percent of Cohort 1 grantees satisfied the criteria for an active IGA per the DIB's business metric by providing sufficient supporting details for a recent transaction; this percentage can be viewed as a medium-term version of the business metric (Appendix Table A8).[11] In contrast, more than 98 percent of Cohort 1 grantees had an active IGA according to the IGA survey conducted a little more than one year earlier.

There was a modest gender gap in the medium-term version of the business metric: 77 percent of women grantees satisfied the criteria for an active IGA compared to 70 percent of men. The top income quartile also had a business metric that was between 7 and 12 percentage points higher than the bottom two quartiles. Differences in this metric based on age and nationality were smaller, between 4 and 5 percentage points, and favored adults and Jordanians.

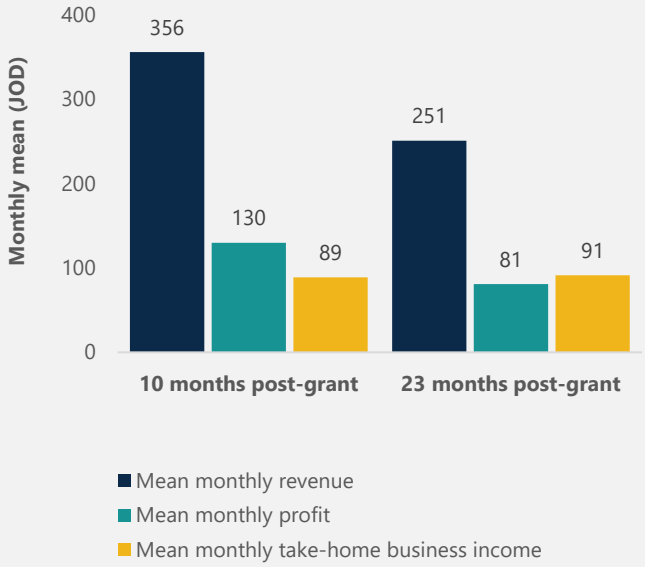
Grantees with active IGAs at the two-year mark were still conducting frequent sales transactions. Of those with active IGAs, almost all continued to meet the business metric criteria through a recent sales transaction rather than other criteria. The mean number of days since the most recent sales transaction was 11 days, which still indicates relatively frequent transactions compared to the business metric criterion of 60 days (Appendix Table A9a). However, only about 50 percent of impact survey respondents who reported a recent sales transaction were able to provide written documentation of that transaction compared to more than 90 percent of IGA survey respondents, which implies that many grantees did not maintain robust business record-keeping practices after post-grant supports like one-to-one mentoring had ceased.

Most grantees' businesses remained profitable about two-years after receiving grants but reported revenues and profits have declined over time. Among active IGAs at the two-year mark, 80 percent reported positive profits in the preceding month (Appendix Table A9a). Mean revenues and profits were both about one-third lower than those reported in the 10-month IGA survey (**Figure VI.1;** Appendix Table A9a). However, with the decline in business-record-keeping it was also much less common for grantees with active IGAs to base their estimates of revenue and costs on written business financial records in the impact survey than in the IGA survey (25 percent versus 75 percent). It is therefore possible that the longer-term estimates are not as accurate as the shorter-term ones.

[11] This medium-term version of the business metric is not used for DIB-related payments, but rather to assess how IGAs evolved over time using consistent criteria.

Take-home business income from grant-supported businesses has held steady over time. Despite the decline in reported profits, grant-supported businesses have continued to provide a steady source of household income. Mean self-reported monthly take-home business income among active Cohort 1 businesses at the two-year mark was 91 JOD ($128, or PPP $476) (**Figure VI.1**, Appendix Table A9a), almost identical to that reported at the 10-month mark.

Figure VI.1. Business financial metrics for active Cohort 1 grant-supported businesses at 10 and 23 months post-grant



Source: Cohort 1 IGA survey (10 months) and impact survey (23 months)

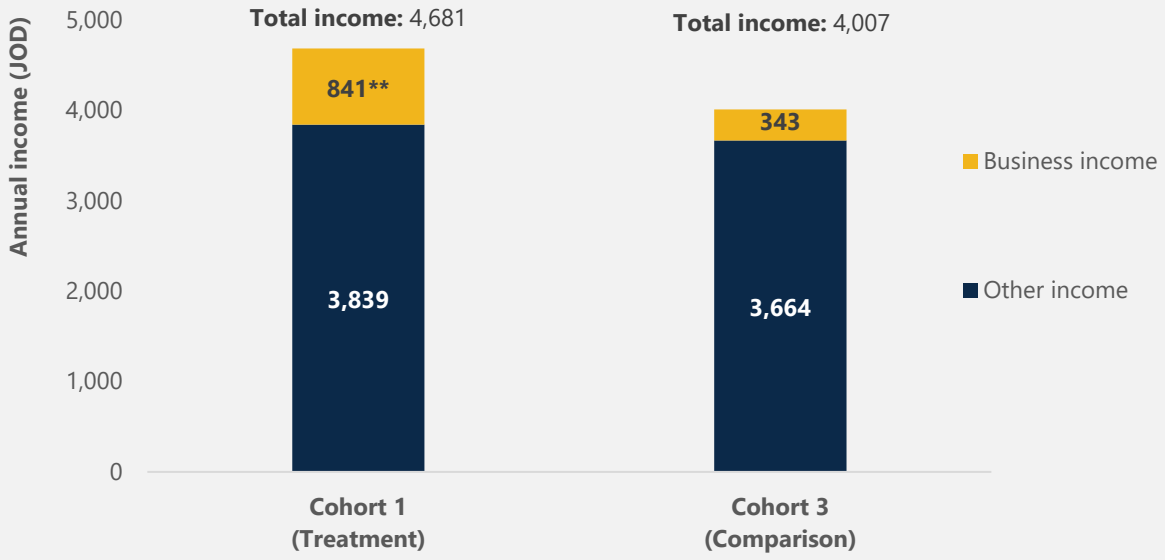
Notes: The samples from the two surveys only partially overlapped, but both were close to representative of the overall population of Cohort 1 grantees.

2. Businesses ownership and household income

Cohort 1 participants reported higher average annual household incomes than matched Cohort 3 participants, primarily driven by their higher incomes from household businesses. Cohort 1 participants—who comprised about two-thirds grantees and one-third non-grantees—were substantially more likely to report that their household owned a business compared to matched Cohort 3 participants (63 versus 35 percent), and those with businesses had more businesses per household on average (1.2 versus 1.0). The relatively high rates of business ownership by matched Cohort 3 participants even though they had not yet received grants likely reflects a mix of (1) existing business operated by the participant, (2) existing businesses operated by other household members, and (3) launch or anticipation of new business activities by the participant in advance of grant receipt.

Despite this higher-than-expected business activity among Cohort 3, Cohort 1 participants reported more than double the mean annual take-home business income reported by matched Cohort 3 participants, a difference of 498 JOD ($702, or PPP $1,853) (**Figure VI.2**, Appendix Table A10). This increased business income explains most of the additional 674 JOD ($950, or PPP $2,507) in total household income reported by Cohort 1 participants relative to matched Cohort 3 participants, a statistically significant difference equivalent to 17 percent or 0.24 standard deviations. Cohort 1 participants also reported higher annual social assistance income and income from pensions, wages, and assets compared to matched Cohort 3 participants, but these differences were modest and not statistically significant except for social assistance.

Figure VI.2. Impacts on annual business income and total household income



Source: Impact survey

Notes: Samples include grantees and non-grantees.

** Difference significantly different from zero at the .01 level, two-tailed test.

B. Reduced poverty of participant households

1. Household consumption

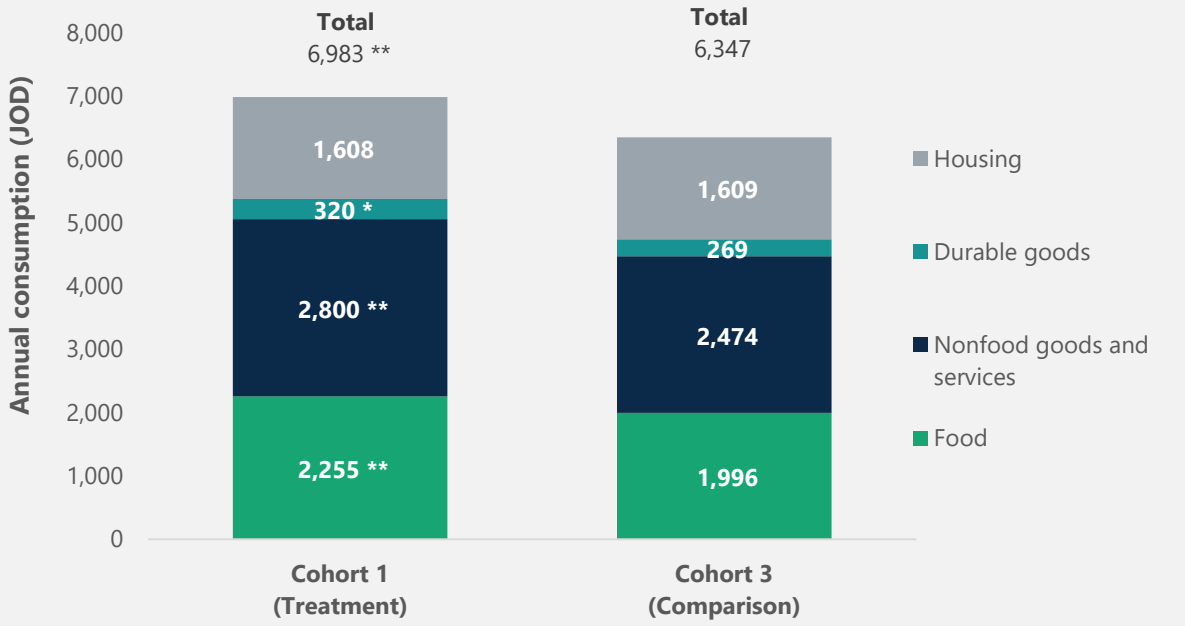
The program led to a 10 percent increase in total household consumption after 23 months, mainly driven by greater food and non-food consumption. The estimated average value of the consumption metric for Cohort 1 households was 6,983 JOD ($9,846, or PPP $25,978) per year, a statistically significant 636 JOD ($897, or PPP $2,366) higher than matched Cohort 3 households (**Figure VI.3**, Appendix Table A11).[12,13] This is equivalent to an impact of 0.225 standard deviations, which just above the threshold at which DIB payments based on the consumption metric were triggered. Examining the four categories of consumption, Cohort 1 households reported higher annual consumption of food (by 259 JOD), non-food

[12] Estimated total annual household consumption is substantially higher than estimated household income (for Cohort 3, 6,347 JOD versus 4,002 JOD). This is largely because the consumption metric includes non-expenditure items, specifically the estimated value of consumption of durable goods as well as the estimated value of housing that is owned or used for free. Using a proxy measure for expenditure, which includes only the value of food consumption, non-food expenditures, and direct spending on housing, total household consumption for Cohort 3 decreases to 5,498 JOD—closer to total household income. Additionally, income is commonly underreported in low and middle-income country contexts, especially among vulnerable populations who rely more on informal and seasonal work and may be reluctant to report income due to concerns about privacy, taxes, and eligibility for social protection programs (Deaton and Grosh 2000). This was a key reason for measuring consumption rather than income as a DIB metric.

[13] Estimated total annual household income is similar to that reported in a recent nationally representative survey of Syrian refugee populations in Jordan (2,983 JOD for Cohort 3 refugees, versus 3,336 JOD in the national survey) (United Nations High Commission for Refugees 2023). Our proxy measure for expenditures is modestly higher than expenditures reported in that survey (4,925 JOD for Cohort 3 refugees, versus 3,852 JOD in the national survey). We view these differences as plausible given differences in samples and measure definitions.

goods and services (by 326 JOD), and durable goods (by 51 JOD) than matched Cohort 3 households, on average. However, the value of housing consumption was almost identical across Cohort 1 and Cohort 3. On a per-person basis, the impact was 116 JOD ($164, or PPP $432) per year, equivalent to 6 percent or 0.13 standard deviations.[14]

Figure VI.3. Impacts on annual household consumption, overall and by category



Source: Impact survey

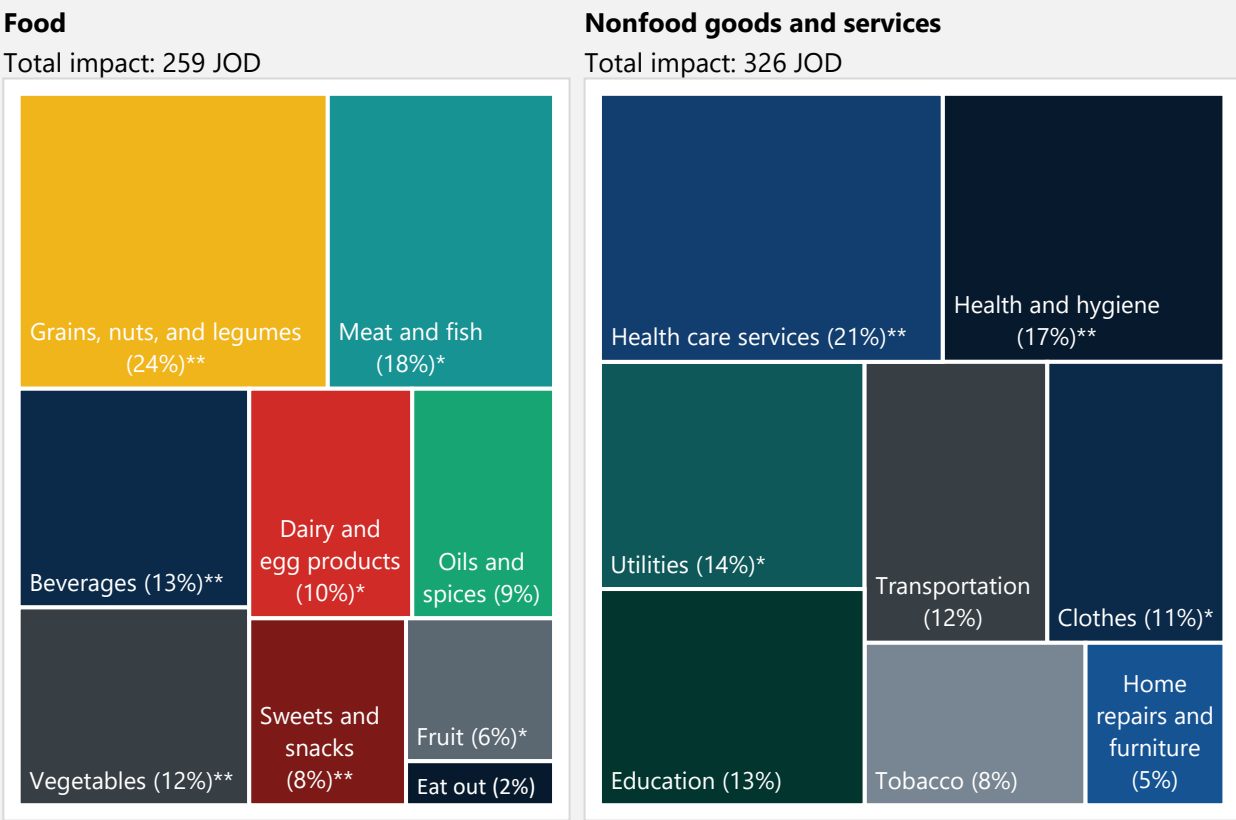
Notes: Samples include grantees and non-grantees.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test

Households were using most of their increased income to increase consumption of nutritious and staple foods, increase their use of health care services, and meet other basic needs like clothing and utilities. Cohort 1 households had higher consumption of all food and non-food categories than matched Cohort 3 households, although not all differences were statistically significant. **Figure VI.4** breaks down the impacts on food and nonfood items and depicts the approximate percentage of the total impact that is associated with each subcategory, as well as whether the impacts were statistically significant. In the food category, Cohort 1 households had statistically significantly higher annual consumption of grains, nuts, and legumes; meat and fish; beverages like coffee, tea, juice, and bottled water; vegetables; and dairy and egg products. In the nonfood category, Cohort 1 households also reported higher average annual expenditures on health care services; health and hygiene products and services (for example, soap, detergent, toothpaste, and haircuts and other salon services); utilities (including electricity, gas, and internet and telephone bills); and clothing.

[14] We used the "adult equivalent" approach, which accounts for differences in consumption between adults and children and economies of scale when estimating per-person consumption rather than simply dividing by the number of household members.

Figure VI.4. Disaggregation of impacts on annual household food and nonfood consumption



Source: Impact survey

Notes: Figure shows the approximate percentage of the total impact that is associated with each subcategory and whether the impacts on each subcategory were statistically significant. Samples include grantees and non-grantees. Totals may not add up to 100 percent due to rounding.

*/** Regression-adjusted differences significantly different from zero at the .05/.01 levels, two-tailed test.

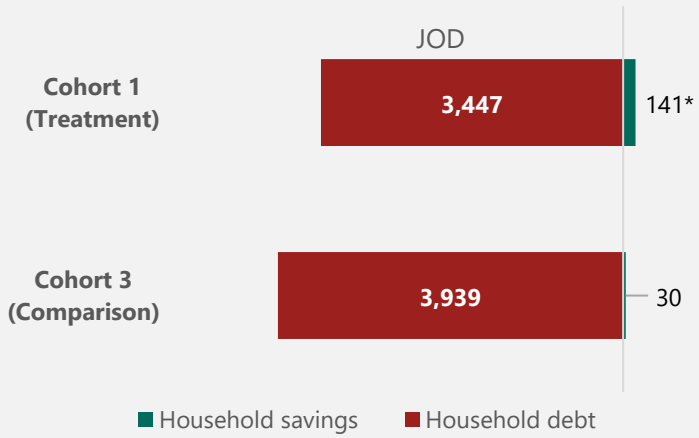
Households were also spending their increased income on increasing their household assets, primarily appliances and electronics. We recalculated the baseline asset index that we used for statistical matching to reflect assets at the time of the impact survey.[15] This index combines several housing characteristics and durable goods ownership and serves as a proxy for household wealth. We find positive impacts of 0.13 standard deviations on this index, driven primarily by household appliances and electronics (Appendix Table A18); this is a modest difference that is equivalent to a gain of about 5 percentile points (von Hippel, 2024). More Cohort 1 households reported owning goods like vacuum cleaners, fans, irons, and freezers than did matched Cohort 3 households (not shown).

[15] Appendix C provides information on the calculation of this index.

2. Debt and savings

Cohort 1 households had high debt levels but few savings, on average; both savings and debt were concentrated in a small number of households. To further assess program households' financial status at the two-year mark, we examined self-reported debt and savings levels. Fewer than 10 percent of households in Cohort 1 and Cohort 3 reported that they had any savings at the time of the survey. As a result, mean savings were modest, although they were higher for Cohort 1 (**Figure VI.5**).

Figure VI.5. Household debt and savings



Source: Impact survey
Notes: Samples include grantees and non-grantees.
***** Regression-adjusted differences significantly different from zero at the .01 level, two-tailed test.

About 85 percent of households in both Cohort 1 and Cohort 3 reported having debts at the time of the impact survey, primarily from formal creditors, relatives, and friends. Mean household debt was 3,447 JOD ($4,860, or PPP $12,823) for Cohort 1 households; this was 12 percent lower than for matched Cohort 3 households, but the difference was not statistically significant because of the high variability in debt (**Figure VI.5**). Median debt levels were much lower, only about 850 JOD ($1.199, or PPP $3,162) in Cohort 1 and 900 JOD ($1,269, or PPP $3,348) in Cohort 3. This indicates that high mean debt levels were driven by a relatively small fraction of households with heavy debt loads. Most of these heavily indebted households were Jordanian rather than refugees and owed large amounts to formal creditors.

3. Food security

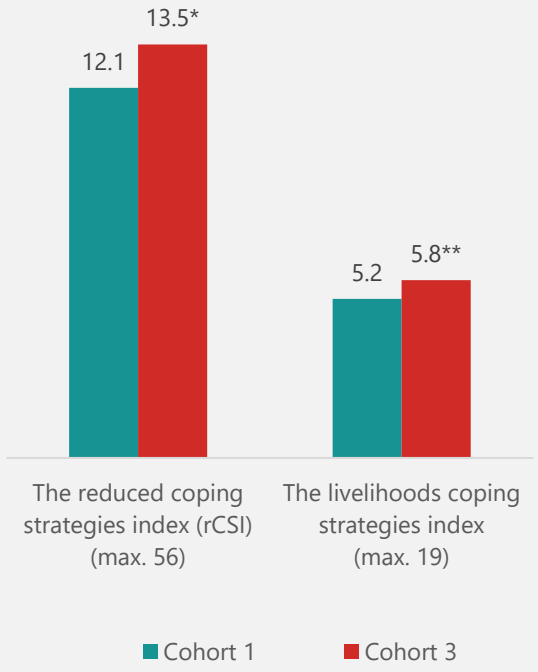
Cohort 1 households reported modestly reduced food insecurity and use of harmful livelihoods coping strategies compared to matched Cohort 3 households. To measure whether increased consumption has translated into reduced use of harmful coping strategies, the impact survey included two scales that have been used in previous studies among refugees in Jordan (REACH 2020). The first, the reduced coping strategy index (rCSI), is a measure of food insecurity that assesses the frequency of harmful food-related strategies undertaken by households to manage food shortages over the previous 7 days (for example, limiting the number or size of meals). The livelihoods coping strategies index examines the broader harmful livelihoods-related strategies that a household has implemented to make ends meet over the previous 30 days (for example, spending savings, selling assets, engaging in high-risk work, or child labor). Both indices categorize strategies in terms of their relative severity based on the local context and culture and assigns more weight to more severe strategies when estimating the index.[16]

[16] For example, in the livelihood based coping strategies index implemented in this evaluation, spending savings is considered a stress-level coping strategy and assigned one point in the index, while taking on jobs that are high-risk, illegal, or socially degrading is considered an emergency-level strategy and is assigned three points in the index.

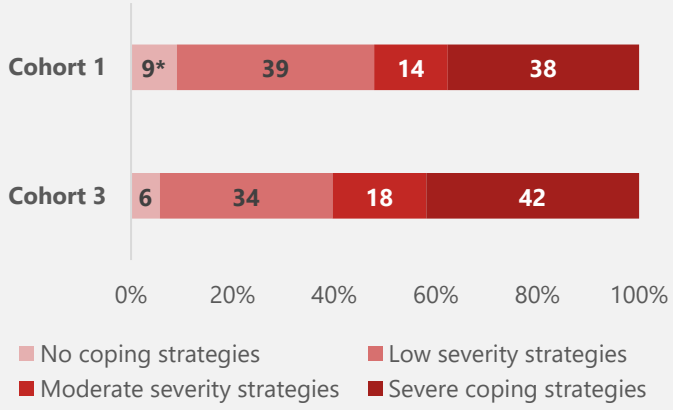
The impact survey shows modest but statistically significant reductions in both overall coping strategies indices, meaning that Cohort 1 households were using fewer negative food- and livelihoods-related coping strategies and using them less frequently at the two-year mark than matched Cohort 3 households (**Figure VI.6**; Appendix Table A12). Despite these positive impacts, it was still common for Cohort 1 households to use relatively severe strategies, suggesting that most were still not able to ensure food security and fully meet their basic needs.

Figure VI.6. Impacts on coping strategies

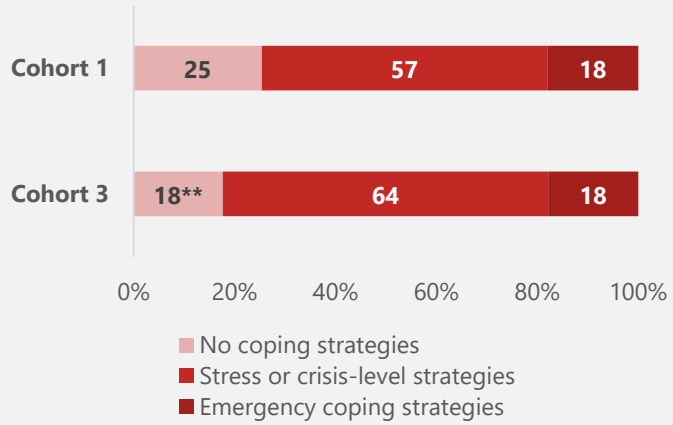
Panel A. Impacts on coping strategies indices



Panel B. Reported use of rCSI (food-related) coping strategies in the past 7 days, by level of severity



Panel C. Reported use of livelihoods coping strategies in the past 30 days, by level of severity



Source: Impact survey

Notes: Samples include grantees and non-grantees. The categorization and weighting of coping strategies is based on REACH (2020).

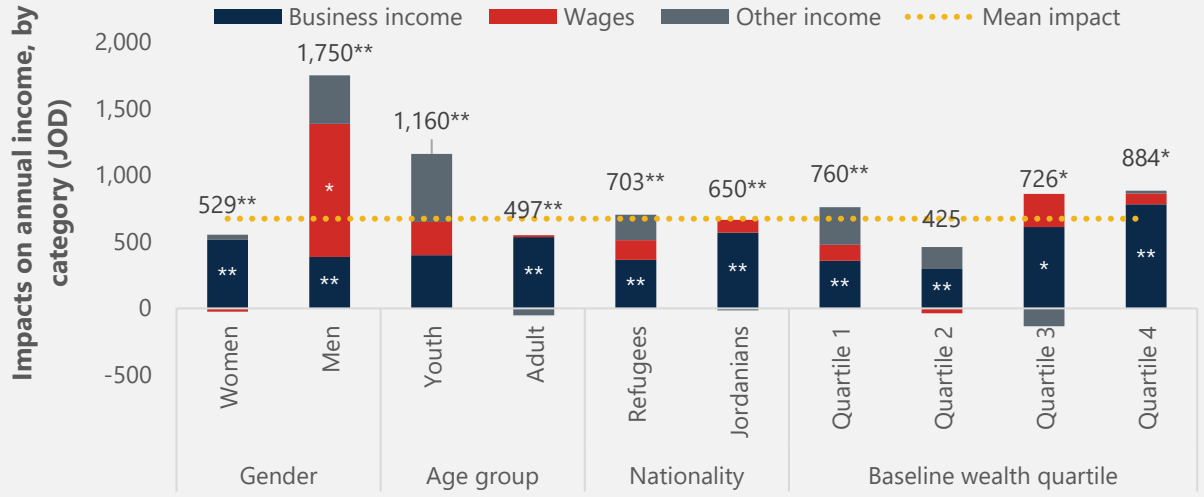
*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test.

4. Differential impacts on economic well-being

Differences in impacts across subgroups likely reflect a complex interplay of demographic, socio-economic, and other household characteristics. Although we report findings for distinct subgroups below, these characteristics intersect in complex ways. For example, more than 90 percent of refugees in the analysis sample were in the bottom two asset quartiles, compared to a little more than 30 percent of Jordanians; 26 percent of youth participants were men compared to 8 percent of adult grantees. It is not possible to fully disentangle the effects of gender, nationality, the age of the participant, household wealth, or other interrelated characteristics because of limited sample sizes and because these factors likely interact with each other in complex ways. As a result of this complexity, and because the consumption-focused impact survey did not exhaustively capture all expenditure categories, we are unable to fully explain differences in impacts on economic well-being across subgroups. Nevertheless, below we explore subgroup differences to better describe how the program's impacts were distributed.

Impact on household income were substantially higher for male and youth participants, in part due to higher income from non-business sources. The impacts on self-reported household income for men were more than three times those for women, and the impacts for youth more than double those for adults (**Figure VI.7**, Appendix Table A13). Although most of the impacts on income for the full sample were driven by business income, male and youth participants in Cohort 1 also reported higher income from household members' wages than matched Cohort 3 participants (Appendix Table A14). Households in the highest baseline asset index quartile experienced the largest impacts on income, although there was no clear trend in the impacts across the other quartiles. Differences in impacts on total household income between refugees and Jordanians were modest.

Figure VI.7. Impacts on annual household income, by income source and subgroup.



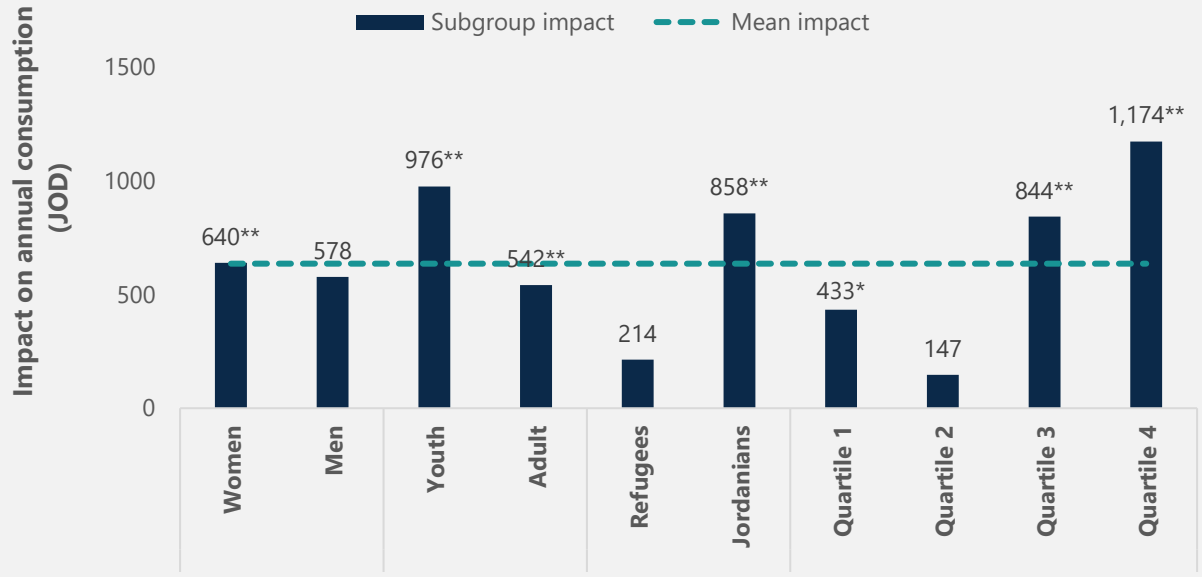
Source: Impact survey

Note: Business income includes all household businesses. Wages include income for all household members including both informal and formal employment,). Other income includes social assistance programs, income from assets and pensions, remittances, and support from family and neighbors.

*/** The subgroup-specific impact is significantly different from zero at the .05/.01 levels, two-tailed test. Additional analysis shows that the differences in impacts between men and women and between youth and adult are statistically significant at the 0.05 level but differences between youth and adults and across baseline wealth quartiles are not.

Impacts on annual household consumption were largest for youth, Jordanians, and households in the highest baseline asset quartile. Larger impacts on income did not consistently result in larger impacts on consumption across subgroups (**Figure VI.8,** Appendix Table A15). Youth benefited from both larger impacts on income and consumption. It appears that these participants are spending their increased income to support their households' basic needs and increase their quality of life, leading to substantial increases in consumption, as well as assets and food security (Appendix Tables A16-A18). Also consistent with the pattern of impacts on income, households that were relatively best off at baseline experienced the largest impacts on consumption. In contrast, while impacts on income for men were much larger than for women, there were no differences in impacts on consumption. Further, refugees and Jordanians had similar impacts on income, but the former had much smaller impacts on consumption. Finally, households who were worst off at baseline experienced moderate impacts on income but small impacts on consumption. The use of income for unmeasured expenditure categories such as financial support to other households (including remittances by refugees to family members still in Syria) or payments towards debts might play a role in explaining these gaps between impacts on income and consumption.[17]

Figure VI.8. Impacts on annual household consumption, by subgroup.



Source: Impact survey

Note: Additional analysis shows that the differences in impacts between groups are only statistically significant for refugees versus Jordanians.

*/** The subgroup-specific impact is significantly different from zero at the .05/.01 levels, two-tailed test.

[17] We are unable to explain these using available data, as the impact survey was designed to measure consumption and not expenditure (beyond expenditure on non-food items and services that were part of the consumption metric).

C. Improved sense of well-being

There is little evidence of impact on life satisfaction or self-confidence after two years. As discussed in Section V, in FGDs conducted about one year after grants were disbursed, several grantees indicated that their sense of self-confidence increased as a result of participating in the program and the increased financial independence that followed. Grantees who participated in FGDs also generally reported a positive outlook towards the future, expressing a stronger sense of motivation and resilience.

In the impact survey, we measured two internationally validated indices to quantitatively assess impacts on life satisfaction and self-confidence on Cohort 1 participants, including both grantees and non-grantees. These indices involved asking respondents to rate a series of statements about their satisfaction with their lives and positive and negative feelings towards themselves. Both indices were nearly equivalent between Cohort 1 participants and matched Cohort 3 participants (**Figure VI.8**, Appendix Table A19), providing little evidence of positive impacts. These findings might be affected by the impact survey having been conducted after Cohort 3 was selected for the program, and in most cases after they had completed the business training. This might have increased their life satisfaction and self-confidence, as they would have been looking forward to potentially receiving a grant and implementing their business plans .

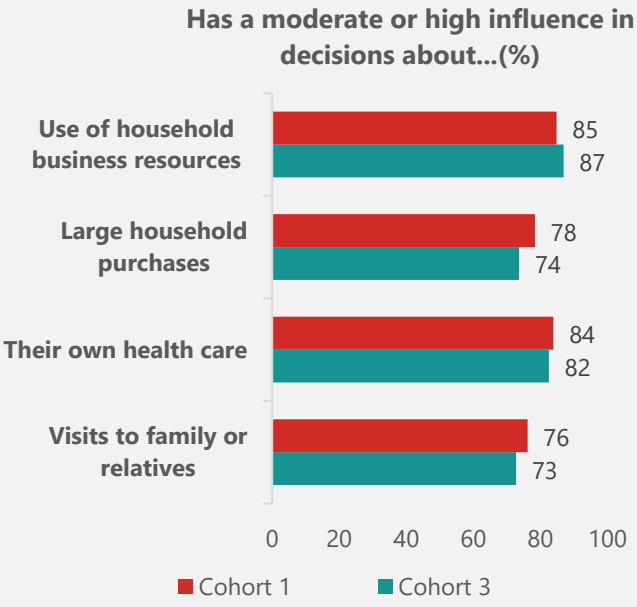
We found no impacts on women's influence on household decision making or their freedom of movement. Our measure of household decision-making focused on respondents' perceived extent of influence in decisions. In both cohorts, the vast majority of female program participants whose households had an active business reported that they had moderate or high influence in deciding how household business resources are used (**Figure VI.9**, Appendix Table A19). Although there were no substantive differences in this outcome between matched female participants in Cohorts 1 and 3, its

Figure VI.8. Satisfaction with life and self-confidence scales

Source: Impact survey

Notes: Samples include grantees and non-grantees. These outcomes were measured using a modified version of the Diener satisfaction with life scale (Diener (1985), and the self-confidence scale was measured using the Rosenberg self-confidence scale (Rosenberg 1965). Neither difference is statistically significant.

Figure VI.9. Female participants' perceived influence over household decision-making



Source: Impact survey

Notes: Samples include grantees and non-grantees. Use of household business resources is restricted only to those women who reported that their households had an active business. No differences were statistically significant.

high rate in both cohorts is still a positive finding because it suggests that female participants can control the income they earn from their small business project. The percentage of female program participants who reported playing an influential role in deciding on large household purchases, their own health care, and visits to family or relatives was also high (about 7 or 8 in 10 respondents) and similar in Cohorts 1 and 3. We also measured a freedom of movement index, based on questions about respondents' ability to visit places outside the home independently of permission from (or accompaniment by) male household members. The mean index was 2.6 out of 4 for Cohort 1 and identical for matched Cohort 3 participants.

Beyond women's participation in decision-making and freedom of movement, the qualitative data suggest that the program may have led to positive normative change around gender in other ways. For example, the establishment of new household businesses created new opportunities for men and women to collaborate economically; both male and female grantees reported working with their spouse to calculate profits, advertise products, and share skills learned during trainings. Some female participants in the FGDs explained that they had husbands or brothers who did not initially support their business ventures, but their attitudes shifted after seeing their female family member succeed. The project may also have had spillover benefits onto other girls and women, since many female participants now see themselves as role models in their communities, and several have encouraged female neighbors or family members to start their own businesses.

> "You start feeling like you're making a positive contribution to society and becoming a role model to your children."
>
> Female grantee, Cohort 1, Amman

D. Improved welfare of refugees and host community

Improved welfare of refugees and host community is considered a longer-term outcome that is likely to emerge beyond two years post-grant. In this section we explore early signs of improvement in well-being based on the impact survey and qualitative data collected as part of the process evaluation.

1. Improved living conditions

Participants have invested in household appliances and electronics that can improve their quality of life, but changes in the physical condition of their housing have been limited. As discussed earlier, Cohort 1 households had higher consumption of durable goods and ownership of appliances and electronics than matched Cohort 3 households, which would contribute to a higher standard of living. In contrast, there were few meaningful differences between Cohorts 1 and 3 in the rates of household home ownership, housing materials, the number of people per room, expenditures on home repairs, or the value of housing at the time of the impact survey.

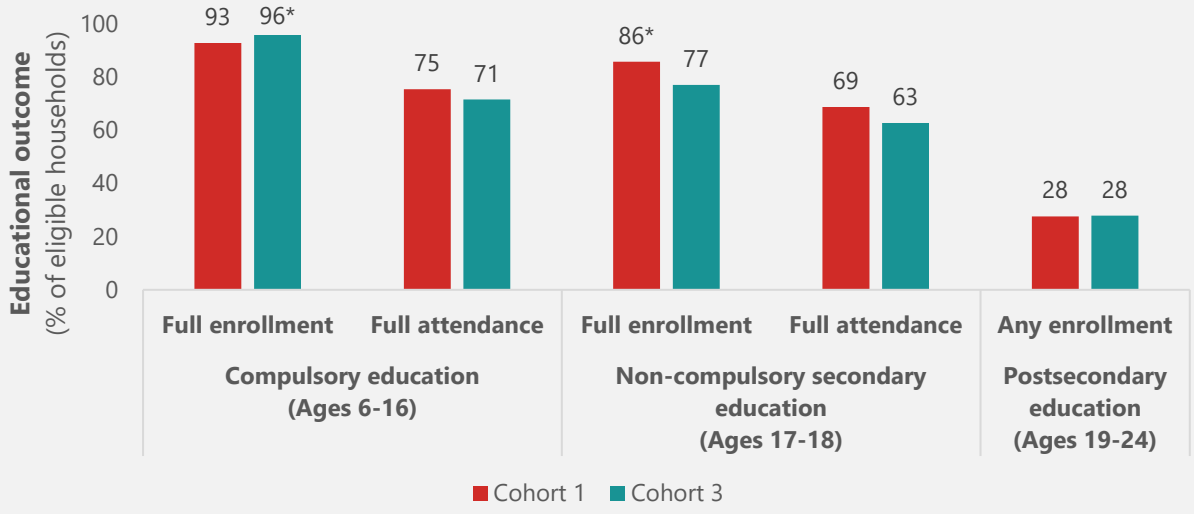
2. Educational outcomes

Impacts on educational outcomes were limited, but there is some evidence of increased enrollment in secondary education. Children from vulnerable households who attend school for longer and learn more will have greater social and human capital that will help them to contribute to their families and communities, while refugees and Jordanians who attend school together may contribute to improved social cohesion over time. Public education in Jordan consists of free and compulsory education from ages 6 to 16, followed non-compulsory but free public education for students aged 17 and 18, and finally by a variety of postsecondary education and training options. To assess changes in educational outcomes,

we measured (1) whether all school-age household members were enrolled in the relevant level of education, (2) among enrolled household members in compulsory or secondary education, whether all attended every day that their school was open during the previous week, and (3) whether any household members aged between 19 and 24 were enrolled in further education or training.

Enrollment in compulsory education was very high in Cohorts 1 and 3, with more than 9 in 10 households with children 6–16 years old enrolling all their children (**Figure V.10**, Appendix Table A20). This measure was slightly lower for Cohort 1, but the difference was small and was concentrated among 16-year-olds. In contrast, Cohort 1 households were 9 percentage points more likely than matched Cohort 3 households to have all members aged 17 or 18 enrolled in school, a statistically significant difference that suggests positive effects when enrollment is discretionary. Effects on school attendance or on enrollment of any young household members aged between 19 and 24 in postsecondary education or training were small and not statistically significant.

Figure VI.10. School enrollment and attendance by age group



Source: Impact survey

Notes: Enrollment statistics do not include 6-year-olds because the survey was conducted close to or after the end of the school year and most children who were 6 years old were likely not eligible for enrollment during the previous school year. They are included in the attendance since this is conditional on having been enrolled. All analyses are conditional on having household members in the relevant age range. Full enrollment means that all children in an age range were enrolled in school while full attendance means that all enrolled students attended on all days that school was open. The postsecondary measure reflects whether at least one household member aged 19 to 24 was enrolled in school.

3. Community integration

Although the program successfully promoted positive interactions between refugees and host community participants, evidence of broader impacts on community integration is limited. FGD respondents highlighted the program's positive effects on their social circles and business networks, with the program increasing the number of acquaintances in their community and building skills and comfort in interacting with them for social and business purposes. Several have established WhatsApp group to socialize, collaborate, or exchange items with other program participants.

Some FGD respondents reported that the program has fostered sustained social connections between participants that transcend national boundaries. For example, respondents in one FGD reported that both Jordanian and Syrian participants still meet together for social activities, and a Syrian respondent said that the program expanded her social network of Jordanians. However, these experiences were not universal. One FGD respondent noted that Syrians and Jordanians interacted well during the trainings, but they did not stay in contact afterwards. Another respondent explained that although Syrians and Jordanians have developed positive social relationships through buying and selling products to each other, members of these groups typically do not work together on business projects or share expertise with each other, resulting in limited collaboration and exchange between the two groups. Our data collection did not focus on broader effects on community integration associated with business activities of program participants (for example, Jordanians' perceptions of refugees in the community), which would likely take longer to emerge and would require a broader study sample beyond program participants.

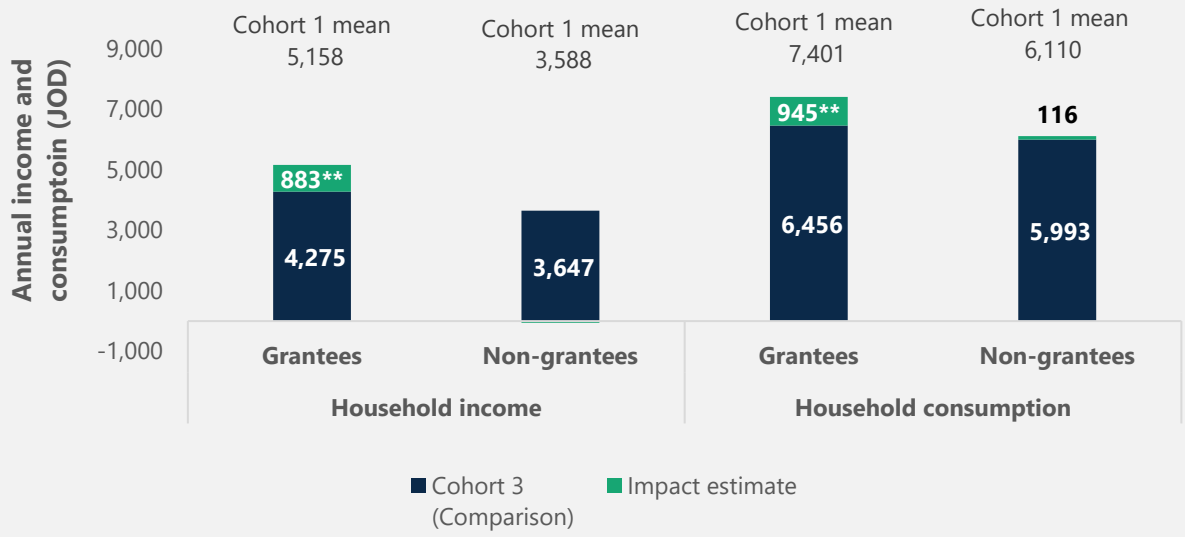
E. Impact estimates by grantee status

We conducted a further analysis where we rematched the sample to include grantee status (for Cohort 3, this was future grantee status as grants had not yet been awarded at the time of the survey), in addition to the socio-demographic characteristics used to match the full sample. This enabled us to estimate impacts separately for Cohort 1 grantees and non-grantees by comparing them to matched Cohort 3 grantees and non-grantees, respectively. These samples are smaller than the overall matched sample and hence provide more limited power to detect statistically significant impacts. Nevertheless, conducting the impact analysis by grantee status provides insights into the extent to which the receipt of grants and post-grant support, rather than the business skills or other trainings, are driving the overall findings .

Impacts on income, consumption and other measures of economic well-being were driven almost entirely by grantees, with near-zero impacts among non-grantees. The impact on household income for Cohort 1 grantees two years following grant receipt was 883 JOD ($1,245, or PPP $3,285), equivalent to a 21 percent or 0.32 standard deviations increase relative to matched Cohort 3 future grantees (**Figure VI.11**). The impact on annual household consumption for Cohort 1 grantees at the two-year mark was 945 JOD ($1,332, or PPP $3,515), equivalent to a 15 percent or 0.36 standard deviation increase relative to matched Cohort 3 grantees. These impacts on income and consumption for grantees are about one-third and one-half greater, respectively, than the impacts for the full matched sample discussed previously.[18] In contrast, the impact for non-grantees on income and consumption were both close to zero. Detailed findings for these and additional outcomes are provided in Appendix Tables A21 and A22.

[18] The comparison with the previously reported impacts for the full matched sample is not strictly correct because adding grantee status as a matching variable led to some sample size loss. A stricter comparison would be to the impact for the full sample of grantees and non-grantees after rematching on grantee status, which is 732 JOD ($1,032, or $2,723). However, this does not materially affect our conclusion that the impact for grantees was substantially larger than the overall impact.

Figure VI.11. Impacts on measures of economic well-being, by grantee status



Source: Impact survey

Note: Impacts on non-grantees' income were -60 JOD, and not statistically significant.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test

VII. Conclusion

In this concluding section, we summarize the findings for each research question, contextualize the impact findings in the broader literature of similar programs, and discuss lessons and recommendations for future programming.

A. Summary of findings

The evaluation provides strong evidence to support the program logic. It was implemented effectively, improved upon that implementation over time, and exceeded targets for establishing active IGAs. Moreover, it had positive effects on key medium and long-term well-being outcomes, especially among those who received cash grants. In **Table VII.1**, we summarize findings for the study's research questions.

Table VII.1. Summary of findings, by research question.

Research question	Summary of findings
What percentage of grant recipients were actively engaged in IGAs 10 months after grant disbursement?	• 98 percent of grantees had an active IGA 10 months after grant disbursement.
What were the impacts of program participation on social and economic wellbeing 24 months after grant disbursement?	• Nearly two years after grant disbursement, 76 percent of Cohort 1 grantees still had an active IGA. • Increased business ownership and income translated into positive impacts on total annual household income, which in turn resulted in positive impacts on annual household consumption of 636 JOD ($897, or PPP $2,366), or 0.22 standard deviations. • The program also had modest, positive impacts on reported household savings (although savings were still uncommon) and durable household assets. • We find no evidence of impacts on self-confidence, satisfaction with life, or women's social or economic empowerment, although the impact study design was not well-suited to identifying some of these impacts and there is some qualitative evidence of changes. • The program significantly reduced food insecurity and the use of harmful coping strategies. There is also evidence that the program modestly increased enrollment in non-compulsory secondary education. Evidence of changes in the use of social assistance were inconclusive. • Impacts on household consumption were largest for youth, Jordanians, and households that were the best off at baseline, although intersectionality across subgroups and the interplay with impacts on income are complex.
What were the key elements of the program that led to achieving the desired program outcomes?	• Participants viewed the business skills training as being well-delivered and very valuable. Participants' feedback on life skills and technical trainings was more mixed. • The program used a rigorous grant selection process to identify businesses with strong potential for success. Grants provided a strong foundation for participants' businesses and were critical to impacts on well-being, as impacts on non-grantees were negligible. • Grantees also benefitted substantially from post-grant supports, with one-to-one mentorship being especially valuable.

Research question	Summary of findings
What is the community and business environment in which participants live and work?	• Key facilitators to business survival and growth were family, community, and program-provided post-grant support, while key barriers included rising costs, competition, and challenges with business registration. Refugees and women face additional barriers and constraints to operating and growing their businesses. • The program helped participants expand their social circles and business networks, but there was limited evidence of longer-term changes in community integration between Jordanians and refugees in the timeframe of the evaluation. • There is mixed evidence as to program effects on the use of social protection schemes. Some participants became more involved with support services through Siraj centers, but others were less likely to utilize these services as a result of becoming more self-reliant.

B. Comparison to impacts found in benchmark studies and other similar programs

To contextualize the impact findings, we compared them to those of studies of related programs in the literature.

The impacts on the consumption metric and household income are near the upper range of impacts found in the reference studies that were used to set the thresholds for the DIB payments. These studies from Sub-Saharan Africa and South Asia, which were used for benchmarking because they had both a cash grant and training component, had impacts of between 0.07 and 0.38 standard deviations on consumption or expenditure (Keaveney et al., 2018, Appendix F). In comparison, our estimated impacts were 0.22 standard deviations for consumption for all participants, and 0.36 standard deviations for grantees only; the latter may be a more relevant comparison given that all participants in the reference studies typically received the full program support, including cash. Impacts on income in the benchmarking studies were between 0.12 and 0.30 standard deviation, compared to our estimated impacts of 0.24 standard deviations for income for all participants, and 0.32 standard deviations for grantees only.

Impacts were also similar to impacts of livelihoods and cash transfer treatments from the Middle East and North Africa (MENA) region. We find the following relevant comparisons to our estimated impacts, which in percentage terms were 10 percent for consumption and 16 percent for income for all participants, and 15 percent for consumption and 21 percent for income for grantees only. Again, the grantee-only comparisons may be more relevant given that other programs typically provided full support to all participants.

- **A cash-for-work program** implemented in a Jordanian refugee camp increased income by 23 percent after 18 months of program implementation, but only among semi-skilled workers (Lombardini and Mager 2019). This program, which primarily served men, positively affected one measure of self-esteem focused on whether they were making a positive contribution to their family, but not the three other measures. It had mixed impacts on men's perspectives about gender roles.
- Two impact evaluations of World Food Program **unconditional cash transfer programs** that provided Syrian refugees in Lebanon with regular transfers found positive effects on economic well-being and other outcomes, but not long-lasting benefits. Chaaban et al. (2020) found that the program increased

total household expenditures by 20 percent, including significant increases in food expenditures, immediately after receiving the transfers for 16 to 22 months. It also had positive impacts on food security, enrollment, access to health care and self-reported mental health. However, the impacts of the transfers faded several months after the withdrawal of support. Similarly, Altindağ and O'Connell (2023) found a 19 percent increase in expenditures for eligible households during the period of support, and that higher expenditures were used to meet basic needs. The transfers also increased savings and reduced the use of livelihood coping strategies. However, impacts were not sustained once support was discontinued.

- A randomized controlled trial of a program in Jordan that provided up to **6 months of wage subsidies** to recent female community college graduates (who are likely different from our study population in many ways) showed initial positive effects on employment and income (17 percent among employed respondents) that faded once the voucher program ended (Groh et. al. 2012). Additional study arms that provided training or both training and vouchers did not show any significant effects.
- A **cash transfer scheme** targeting poor and vulnerable households in Egypt reduced debt, increased assets by around 0.18 standard deviations relative to baseline and increased primary school enrollment by 9 percentage points after 15 months of receiving the transfers (El Enbaby et al. 2022; El Enbaby et al. 2024). However, it had no or mixed effects on poverty, consumption, nutrition, mental health, or on women's empowerment outcomes.

In general, we consider our findings to be within or above the range of estimated impacts on economic well-being, based on the limited existing literature for related programs outside the MENA region and for livelihoods and cash transfer programs in the region. Together, these studies indicate that kickstarting self-reliance through IGAs can generate long-term impacts on expenditure and consumption that are equivalent to providing ongoing regular cash support. However, even when changes in economic well-being occur, it remains challenging to measurably shift non-economic outcomes like women's empowerment, self-confidence, or mental health through social protection or economic empowerment programs alone.

C. Cost-effectiveness

We estimated two different benefit streams based on the findings of the evaluation: (1) business profits, measured through the IGA validation study and impact evaluation, and (2) household consumption, measured through the impact evaluation. Since the evaluation findings suggest that business profits are closely related to increased household consumption through increased take-home business income, these should be considered alternative, overlapping, approaches to measuring program benefits, and should not be added together to avoid double-counting. Appendix G provides additional details on methods, and findings for this analysis.

We find that the program likely provides substantial returns on investment. Over 10 years, we estimate that the project would generate $20.1 million in business profits or $22.0 million in additional household consumption. This translates into $9.9 million ($2,900 per grantee) in net business profits or $11.8 million (3,500 per grantee) in increased consumption after subtracting program costs. The benefit-cost ratio for business profits was $1.98, meaning that the program generated $1.98 in net consumption for every dollar invested; the benefit-cost ratio for household consumption is 2.16. Although these

estimates rely on assumptions about the long-term viability and growth of program-funded businesses, a further analysis showed that the program is cost-effective under a wide variety of assumptions.

The program was also cost-effective compared to other, similar programs. We compared our cost-effectiveness findings to a review by Sulaiman et al. (2016) of cost-effectiveness analyses of nearly 50 social protection programs implemented in low- and middle-income countries. The review includes only studies that measured impacts on household consumption and/or income, which enables direct comparisons to our findings on cost-effectiveness in terms of consumption. We find that the program was cost-effective compared to three different kinds of programs, each of which has some components in common with the DIB program: cash transfers, livelihoods programs, and ultra-poor graduation programs. At a cost of nearly $3,000 per grantee, the DIB program had a higher per-beneficiary cost than the average across livelihoods programs ($1,147), cash transfer programs ($232 plus unreported administrative costs), and graduation programs ($1,148). Nevertheless, given the large impacts of the DIB program on grantees' consumption, a simplified benefit-cost ratio for the program was more than double that for the typical livelihoods program, two-thirds higher than that for the typical cash transfer program, and more than four times that for the typical graduation program.[19]

D. Lessons learned and recommendations

The experience of the DIB program suggests several lessons and recommendations for future programing:

The multi-year flexible funding provided by the DIB, its use of both short- and longer-term payment metrics, and multiple stages of measurement, helped to align implementer incentives with program objectives. The Refugee Livelihoods DIB was structured as a multi-year agreement that guaranteed funding for three program cohorts subject to satisfactory performance for earlier cohorts, with NEF given substantial flexibility in how to implement the program and spend these funds. The payment metrics for the DIB included both a short-term outcome (the business metric) measured across multiple cohorts and one longer-term outcome (the consumption metric) based on an impact evaluation for Cohort 1. The business metric comprised the bulk of payments, given that it was more directly in the program's control and had targets that were set based on a long history of similar programs. DIB parties also used it to assess whether the program's performance at intermediate points was satisfactory. In contrast, the consumption metric had a smaller payment and was treated as a "bonus," which reflected the greater uncertainty around the program's likely effects on this longer-term outcome—especially given limited studies of comparable programs to set DIB benchmarks for this metric.

The DIB design had several positive effects on program implementation that are supported by interviewers with program implementers, survey findings across cohorts, and the experiences of the evaluation team. First, the guaranteed funding and programmatic and budgetary flexibility offered by the DIB funding model has encouraged NEF to test multiple activities and approaches, collect and analyze

[19] Specifically, for comparability to the estimates in Sulaiman et al. (2016), we computed a simplified benefit-cost ratio as the point-in-time impact on annual consumption (that is, at the time of the impact survey, about two years after grants were disbursed) divided by total program costs per grantee. This ratio is 0.49 for the DIB program, compared to averages of 0.20 for livelihoods programs, 0.29 for cash transfer programs, and 0.20 for graduation programs reported in Sulaiman et al. (2016). Although the point in time at which impacts were measured differs across studies, this simple ratio offers a ready metric for comparing cost effectiveness across numerous studies of related programs.

data at each phase to reflect on their effectiveness, and improve their approaches over time. Their internal monitoring processes, combined with feedback from the external evaluation team, have resulted in measurable improvements in implementation across cohorts. Second, the two DIB payment metrics struck a good balance between balancing DIB parties' financial risk and sufficiently incentivizing outcomes. The short-term business metric incentivized the program to carefully select grantees and to provide them with the support they need to maintain their businesses over time, which helped ensure that the key pathway to longer-term impacts in the program logic was realized. The inclusion of the consumption metric incentivized the implementation team to maximize the long-term sustainability of the businesses that participants established. At the same time, having this metric as a lower stakes "bonus" measure avoided introducing unreasonable financial risk to the DIB parties given the uncertainty describe above. Third, although DIB payments will only be made at the end of the program, the multi-cohort approach and associated multi-step evaluation with several intermediate measurement and reporting stages has fostered a collaborative, mutually supportive relationship between NEF and the independent evaluation team. This has supported ongoing improvements in NEF's implementation and data quality.

Expenditures may be more suitable for use as a measure of economic well-being and a DIB payment metric in this context than consumption. At the DIB design stage, household consumption was correctly highlighted as the theoretically preferred measure of well-being, given challenges associated with accurately and reliably measuring income in low- and middle-income countries (Keaveney et al. 2018). However, in retrospect, we believe that household expenditures would have served as a more practical but still fit-for-purpose DIB payment metric, for several reasons. First, although we benefitted from a data collection team that was experienced measuring consumption, it was often still challenging for respondents to accurately recall details of food consumption by all household members over the past week. Given that these vulnerable households typically purchase the food they consume on a weekly or daily basis, we would expect expenditure on food items to be highly correlated with consumption, while being easier to report. Indeed, in practice it was common for respondents to think through consumption of many food items in terms of expenditures. Consumption of non-food items and services was in any case measured as expenditures over the relevant reference period, per standard practice. Second, the proper calculation of durables consumption required information about estimated value of durables, which respondents also found challenging to report accurately; measuring expenditure on purchasing durables over a one-year period would have been more straightforward. Third, the standard measure of consumption excludes some categories of expenditure, like debt repayments and remittances, which reflect household economic well-being and may be especially important in this context.[20] Overall, given that the aim of the evaluation was to compare economic well-being between a treatment and comparison group rather than to produce an accurate stand-alone measure of household consumption, an expenditure-based measure might have been preferable.

The use of local CBOs as a hub for services can strengthen implementation effectiveness and sustainability. Qualitative interviews with program implementers highlighted the critical role that CBOs played in the success of implementation. Through their longstanding presence in the community the

[20] As mentioned earlier, the DIB's consumption metric was originally adjusted to account for debt repayments, but the proposed adjustment (which focused on repayment of pre-program debt to avoid double-counting consumption) was not practical to measure.

CBOs helped to support broad-based recruitment efforts and built participant trust and confidence in the program. In turn, the CBO leaders reported greater recognition and trust from the community, a greater awareness of community needs, and a stronger capacity to meet them than before implementing the program. The CBOs also worked with NEF to develop the program, collect participant feedback, review and interpret monitoring data, and use it to inform site-specific and program-wide improvements. As a result of their involvement, they were better able to meet local needs, overcome barriers to participation among women and youth, and to connect program participants with further training and other services. NEF also noted that CBOs increased their annual funding outside of the DIB program through their participation in the program, which may reflect the capacity building that occurred through partnership with NEF and through managing such a complex, long-term investment in their communities. NEF found that CBOs' legal, operational, financial, and technical capacity, as well as their previous partnerships and long-term viability, were all important factors to successful implementation in local communities. These lessons learned from the CBO partnership process are informing a partnership assessment tool that NEF is developing to inform CBO selection and targeted capacity building activities on future projects. Taken together, this project highlights the value of locally led implementation of livelihood programs, with appropriate support and capacity building from larger national or international organizations with the relevant capacity, experience, and local knowledge. This approach can also strengthen pathways to localization by increasing the ability of CBOs to manage and run programs independently of international organizations. The CBO partnerships may have also contributed to greater program cost-effectiveness, since many aspects of the program relied on existing CBO infrastructure, staff, and vendor relationships, potentially reducing administrative burdens relative to a program that needed to build them from scratch. (Appendix Table G.2 provides a summary of roles and responsibilities for CBO versus NEF staff).

While the program model shows promise for adaptation and scaling to other contexts, the findings also suggest that additional, targeted supports may be needed to ensure that the benefits of the program are distributed more equitably. The grant selection process, while designed to be merit-based and support the most promising of business plans, may have inadvertently favored Jordanians, adults, and participants who were somewhat more socio-economically advantaged prior to joining the program. Further, qualitative evidence suggests that women and refugees faced additional barriers with operating their businesses, while subgroup findings show that some subgroups (women, refugees, the economically worst-off at baseline) experienced smaller program impacts on income and/or consumption than others. This suggests that support leading up to grant selection as well as post-program support might need further tailoring to carefully identify and address the context-specific barriers faced by the most vulnerable subgroups. For example, since women report spending less time on their businesses, which reduces their income potential, further expanding access to childcare services during as well as after training may help to promote greater gender equity in program outcomes. More support for transportation and an increased emphasis on digital sales and marketing could also help to overcome cultural and logistical constraints on women's business activities. Similarly, refugees may require strategies and supports to help them overcome legal and financial barriers to business ownership and growth, although the specific strategies and their effectiveness may depend on broader changes in the Jordanian policy context .

The positive findings suggest that the program was both an effective and a cost-effective approach to improving the well-being of participants selected for grants, but it may not be a catch-all

solution for improving the well-being of all vulnerable populations. The findings from the evaluation are largely positive, and the magnitude of impacts on income and consumption compare favorably to other comparable livelihoods and cash grants programs in the MENA region. Further, the program's impacts were large enough to justify its costs. However, these findings do not necessarily suggest that this particular program would achieve similar results for the broader vulnerable population in the region, for two main reasons. First, these findings reflect benefits for a carefully selected subpopulation: vulnerable individuals who have the aspirations and the capacity to be entrepreneurs. NEF conducted a robust recruitment process to identify training participants, and then carefully selected a subset of those trainees to receive cash grants based on the strength of their business plans. It is unlikely that the program would be similarly effective if it were scaled up in a way that involved a less stringent selection process that sought to reach a broader vulnerable population. Second, the program's success was built on NEF's extensive experience with and learnings from implementing similar programs in the MENA region and deep understanding of the cultural context. Adapting this program to other countries or by other implementers would need to carefully account for the local business environment and economy; social, cultural, and gender norms; and implementer experience.

References

Al-Hawarin, Ibrahim, Ragui Assaad, and Ahmed Elsayed. "Migration Shocks and Housing: Evidence From the Syrian Refugee Crisis in Jordan." Economic Research Forum Working Paper Series, no. 1213, 2018.

Al-Yahya, Khalid, and Jennifer Airey. "Small and medium sized enterprises in MENA: leveraging growth finance for sustainable development." Shell Foundation and Citi Foundation, 2013. Available at: https://shellfoundation.org/wp-content/uploads/2024/08/shell_foundation_mena_enterprises.pdf.

Altındağ, Onur, and Stephen D. O'Connell. "The short-lived effects of unconditional cash transfers to refugees." Journal of Development Economics 160, 2023.

Aziz, Sahar F., Joanna Gardner, , Tamara Anaie, and Omar Rana. "Towards Sustainability and Empowerment: Reforming America's Syrian Refugee Policy." Newark, New Jersey: Rutgers Law School Research Paper, January 29, 2019. http://dx.doi.org/10.2139/ssrn.3324792.

Borkum, Evan, Paolo Abarcar, Laura Meyer, and Matt Spitzer. "Jordan Refugee Livelihoods Development Impact Bond Evaluation Framework." Washington, DC: Mathematica, 2022.

Brookings Institution. Brookings Institution Global Impact Bond Database, February 1, 2025. Available at https://www.brookings.edu/articles/social-and-development-impact-bonds-by-the-numbers/.

Campos, Francisco, Michael Frese, Leonardo Iacovone, Hillary C. Johnson, David McKenzie, and Mona Mensmann. "Long-Term and Lasting Impacts of Personal Initiative Training on Entrepreneurial Success." Washington, DC: The World Bank, 2024. Available at: https://openknowledge.worldbank.org/entities/publication/f94c4795-0bba-4a38-b10c-caa35f12e46a.

Carter, E., J. Outes Velarde, O. Paul, J.R. Macdonald, and M. Airoldi (2024). "The Impact Bond Dataset: A Tool to Investigate Socially Motivated Cross-Sector Partnerships." Research Data Journal for the Humanities and Social Sciences, pp. 1-14. Available at https://golab.bsg.ox.ac.uk/knowledge-bank/indigo/impact-bond-dataset-v2/.

Chaaban, Jad, Nisreen Salti, Hala Ghattas, Wael Moussa, Alexandra Irani, Zeina Jamaluddine, and Rima Al-Mokdad. "Multi-purpose cash assistance in Lebanon: Impact evaluation on the well-being of Syrian refugees." Cairo, Egypt: Economic Research Forum, 2020. Available at https://erf.org.eg/publications/multi-purpose-cash-assistance-in-lebanon-impact-evaluation-on-the-well-being-of-syrian-refugees/.

Clemens, Michael, Cindy Huang, Jimmy Graham, and Kate Gough. "Migration Is What You Make It: Seven Policy Decisions That Turned Challenges into Opportunities." Washington, DC: Center for Global Development, 2018.

Culbertson, Shelly, Olga Oliker, Ben Baruch, and Ilana Blum. "Rethinking Coordination of Services to Refugees in Urban Areas: Managing the Crisis in Jordan and Lebanon." Santa Monica, California: RAND, 2016.

Deaton, Angus and Margaret Grosh. "Consumption," in "Designing impact survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study, M. Grosh and P. Glewwe, Editors. 2000. Washington, DC: The World Bank, 2000.

Deaton, Angus, and Salman Zaidi. "Guidelines for Constructing Consumption Aggregates for Welfare Analysis." LSMS Working Paper. No. 135. Washington DC: World Bank, 2002.

Diener, E. D., Robert A. Emmons, Randy J. Larsen, and Sharon Griffin. "The Satisfaction with Life Scale." Journal of Personality Assessment, vol. 49, no. 1, 1985, pp. 71-75.

Donald, Aletheia, Gayatri Koolwal, Jeannie Annan, Kathryn Falb, and Markus Goldstein. "Measuring Women's Agency." World Bank Policy Research Working Paper, no. 8148, 2017.

El Enbaby, Hoda, Dalia Elsabbagh, Daniel Gilligan, Naureen Karachiwalla, Bastien Koch, and Sikandra Kurdi. "Impact evaluation report: Egypt's Takaful Cash Transfer Program: Second round report." Washington, DC International Food Policy Research Institute (IFPRI), 2022.

El-Enbaby, Hoda, Bruce Hollingsworth, Jean-François Maystadt, and Saurabh Singhal. "Cash transfers and mental health in Egypt." Economics & Human Biology 54, 2024.

Fallah, Belal, Caroline Krafft, and Jackline Wahba. 2018. "The Impact of Refugees on Employment and Wages in Jordan." Economic Research Forum Working Paper Series, no. 1189, 2019.

Frese, Michael. "Learning from African entrepreneurship—on the psychological function of entrepreneurial preparedness." *Small Business Economics* 63, no. 4, 2024.

Ghanem, Dalia, Sarojini Hirshleifer, and Karen Ortiz-Becerra. "Testing attrition bias in field experiments." Journal of Human Resources, 2023.

Groh, Matthew, Nandini Krishnan, David J. McKenzie, and Tara Vishwanath. "Soft skills or hard cash? The impact of training and wage subsidy programs on female youth employment in Jordan." World Bank Policy Research Working Paper 6141 (2012). Washington, DC: The World Bank. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2116139.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis*, vol. 15, no. 3, 2007, pp. 199-236.

Iacus, Stefano M., Gary King, and Giuseppe Porro. "Causal inference without balance checking: Coarsened exact matching." *Political Analysis*, vol. 20, no. 1, 2012, pp. 1-24.

International Monetary Fund "Implied PPP Conversion Rate, 2024." Available at https://www.imf.org/external/datamapper/PPPEX@WEO/OEMDC/IRQ.

Keaveney, Erika, Mike Duthie, and Andrea Hur. "Development Impact Bond (DIB) for Improving Livelihood for Syrian Refugees and Vulnerable Populations in Jordan and Lebanon: Evaluation Design Report." Arlington, VA: Social Impact, 2018.

Krafft, Caroline, Susan Razzaz, Caitlyn Keo, and Ragui Assaad. "The Number and Characteristics of Syrians in Jordan: A Multi-Source Analysis." Economic Research Forum Working Paper Series, no. 1288, 2019.

Lee, David S. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies*, vol 76, no. 3, 2009, pp.1071-1102.

Legrain, Philippe, and Andrew Burridge. "Seven Steps to Success: Enabling Refugee Entrepreneurs to Flourish." Sydney, Australia: Cities and Settlement Initiative, 2019.

Lenner, Katharina, and Lewis Turner. "Making Refugees Work? The Politics of Integrating Syrian Refugees into the Labor Market in Jordan." Middle East Critique, vol. 28, no. 1, 2019, pp. 65–95. doi:10.1080/19436149.2018.1462601.

Lombardini, Simone, and Franziska Mager. "Livelihoods in the Za'atari Camp: Impact evaluation of Oxfam's Cash for Work activities in the Za'atari camp (Jordan)." Oxford, UK: , Oxfam, 2019. Available at https://oxfamilibrary.openrepository.com/bitstream/handle/10546/620883/er-livelihoods-za%27atari-camp-cash-for-work-221019-en.pdf.

Malaeb, Bilal, and Jackline Wahba. "Impact of Refugees on Immigrants' Labor Market Outcomes." Economic Research Forum Working Paper Series, no. 1194, 2018. Available at: https://erf.org.eg/app/uploads/2018/05/1194_Final1.pdf.

Martínez, Claudia, Esteban Puentes, and Jaime Ruiz-Tagle. "The Effects of Micro-entrepreneurship Programs on Labor Market Performance: Experimental Evidence from Chile." *American Economic Journal: Applied Economics* 10, no. 2, 2018.

Maxwell, Daniel, and Richard Caldwell. "The Coping Strategies Index Field Methods Manual, Second Edition." Rome, Italy: World Food Programme, 2008.

McCaig, Brian and Nina Pavcnik. "Out with the Old and Unproductive, in with the New and Similarly Unproductive." Growth and Labor Markets in Low Income Countries (GLM|LIC) Working Paper No. 23, 2017.

McKenzie, David. "How effective are active labor market policies in developing countries? a critical review of recent evidence." The World Bank Research Observer 32, no. 2, 2017.

McKenzie, David, and Christopher Woodruff. "Business Practices in Small Firms in Developing Countries." *Management Science*, vol. 63, no. 9, 2016, pp. 2967-2981. https://doi.org/10.1287/mnsc.2016.2492.

REACH 2014. "Livelihoods, Employment and Tensions in Jordanian Communities Hosting Syrian Refugees: Thematic Assessment Report." Geneva, Switzerland: REACH, 2014. Available at: https://www.impact-repository.org/document/reach/f9d117b8/reach_jor_report_employmentandtensionsinjordaniancommunitieshostingsyrianrefugees.pdf.

REACH. JORDAN MULTI-SECTOR NEEDS ASSESSMENT: Vulnerable Out-of-Reach Communities (VOC). Geneva, Switzerland: REACH, 2020. Available at: https://reliefweb.int/report/jordan/jordan-multi-sector-needs-assessment-vulnerable-out-reach-communities-voc-may-2020.

ReliefWeb. "The Jordan Response Plan for the Syria Crisis 2016-2019." 9 August 2017. Available at: https://reliefweb.int/report/jordan/jordan-response-plan-syria-crisis-2016-2019.

Rosenberg, Morris. "Rosenberg Self-Esteem Scale (RSE)." Acceptance and Commitment Therapy. Measures Package, vol. 61, no. 52, 1965, pp. 18.

Sulaiman, Munshi, Nathanael Goldberg, Dean Karlan, and Aude de Montesquiou. "Eliminating Extreme Poverty: Comparing the Cost-effectiveness of Livelihood, Cash transfer, and Graduation Approaches." Washington, DC: CGAP, 2016. Available at: https://poverty-action.org/sites/default/files/publications/Forum-Eliminating-Extreme-Poverty-Dec-2016.pdf.

Stuart, Elizabeth A. "Matching methods for causal inference: A review and a look forward." *Statistical Science: a Review Journal of the Institute of Mathematical Statistics*, vol. 25, no. 1, 2010, pp. 1-21.

Tiltnes, Åge A., Huafeng Zhang and Jon Pedersen. "The Living Conditions of Syrian Refugees in Jordan: Results from the 2017-2018 Survey of Syrian Refugees Inside and Outside Camps." ReliefWeb, 2019. Available at: https://reliefweb.int/sites/reliefweb.int/files/resources/67914.pdf.

United Nations Children's Fund (UNICEF). "Geographic Multidimensional Vulnerability Analysis - Jordan." New York, New York: UNICEF, 2020. Available at: https://www.unicef.org/jordan/Geographic-Multidimensional-Vulnerability-Analysis.

United Nations (UN). "Unpacking Gendered Realities in Displacement: The Status of Syrian Refugee Women in Jordan." Cairo, Egypt: United Nations Women Regional Office for Arab States, 2018. https://arabstates.unwomen.org/en/digital-library/publications/2018/12/unpacking-gendered-realities-in-displacement.

United Nations High Commissioner for Refugees (UNHCR). "Vulnerability Assessment Framework (VAF) population report." Geneva, Switzerland: UNHCR, April 2019. Available at: https://data2.unhcr.org/en/documents/download/68856.

United Nations High Commissioner for Refugees (UNHCR). "Socio-economic situation of refugees in Jordan Q2 2023". Amman, Jordan: UNHCR, 2023. Available at https://data.unhcr.org/en/documents/details/103118.

United Nations High Commissioner for Refugees (UNHCR). "Jordan Operational Data Portal." Geneva, Switzerland: UNHCR, November 2024. Available at: https://data.unhcr.org/en/country/jor. Accessed 6 December 2024.

United States Agency for International Development (USAID). "Assessment of the 2015 USAID Guidelines for Cost-Benefit and Cost-Effectiveness Analysis." 2022. Washington, DC: Integra Government Services International LLC, Prepared for the USAID-Funded Learning, Evaluation and Analysis Project III (LEAP III) Activity.

Vijverberg, Wim PM, and Jonathan Henry Haughton. *Household Enterprises in Vietnam: Survival, Growth, and Living Standards*. Vol. 2773. Washington, DC: World Bank Publications, 2002.

Von Hippel, Paul T. "Multiply by 37 (or Divide by 0.027): A Surprisingly Accurate Rule of Thumb for Converting Effect Sizes From Standard Deviations to Percentile Points." Educational Evaluation and Policy Analysis, 2024.

Wall, Melissa, Madeline Otis Campbell, and Dana Janbek. "Refugees, Information Precarity, and Social Inclusion: The Precarious Communication Practices of Syrians Fleeing War." The Handbook of Diasporas, Media, and Culture, 2019, pp. 503-514. https://doi.org/10.1002/9781119236771.ch34.

Yount, Kathryn M., Kristin E. VanderEnde, Sylvie Dodell, and Yuk Fai Cheong. "Measurement of Women's Agency in Egypt: A National Validation Study." Social Indicators Research, vol. 128, no. 3, 2016, pp. 1171–1192. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5010232/

Zhuang, Juzhong; Liang, Zhihong; Lin, Tun; De Guzman, Franklin (2007): Theory and Practice in the Choice of Social Discount Rate for Cost-Benefit Analysis: A Survey, ERD Working Paper Series, No. 94, Asian Development Bank (ADB), Manila, https://hdl.handle.net/11540/1853

Appendix A: Additional tables

Additional Tables: Section III

Table A1. IGA sample characteristics, all cohorts

Characteristic	All grantees	IGA survey sample grantees
Demographic and household characteristics at time of selection:		
Women (%)	80.9	83.1
Youth (%)	25.7	23.5
Refugees (%)	30.0	30.3
Mean age (years)	34.2	34.9
Head of household (%)	30.1	30.0
Has a disability (%)	4.5	4.7
Literate (%)	98.1	98.2
Mean household size	4.3	4.6
Program site (%):		
Amman	21.8	24.7
Irbid	16.9	16.5
Kufrsoum	20.8	20.5
Russeifa	18.0	17.1
Zarqa	22.5	21.2
Asset index		
Quartile 1	25.0	25.3
Quartile 2	25.0	25.4
Quartile 3	25.0	24.8
Quartile 4	25.0	24.5
Grant characteristics:		
Mean grant amount (JOD, NEF-reported)	588	589
Had a business before training began (%, NEF-reported)	10.3	11.6
Business type (%):		
Home food processing	26.5	27.1
Trade in clothes and shoes, fabrics, accessories, perfumes, and make-up	15.4	13.3
Home sewing, tailoring and repair of clothes, shoes, and leather	12.4	10.9
Grocery and food trade	13.0	15.2
Men's and women's salons, beauty centers, and gyms	8.9	9.2
Other	23.9	24.2
Sample sizes	**3,416**	**1,838**

Source: NEF program activity data, NEF vulnerability assessment, and IGA survey (business type, for IGA survey sample only)

Notes: n.a. = Not applicable. The IGA survey sample is weighted to account for differences in cohort size. The asset index is based on housing characteristics and durable goods ownership from NEF's vulnerability assessment, using a principal component analysis for all three cohorts. Youth are defined as being under age 25 at the beginning of the program.

Table A2a. Representativeness of the Cohort 1 impact analysis sample

Characteristic	All Cohort 1 participants	Cohort 1 impact analysis sample	Difference
Demographic, household, and grant characteristics:			
Women (%)	82.4	88.4	-6.0
Youth (%)	30.1	23.5	7.0
Refugees (%)	30.4	35.0	-4.6
Mean age at the time of program intake (years)	35.1	36.8	-1.7
Head of household (%)	29.6	30.1	-0.5
Has a disability (%)	6.9	7.7	-0.8
Literate (%)	97.5	96.7	0.8
Mean household size at time of program intake	5.3	5.3	0.0
Education level (%)			
Less than secondary education	29.1	31.7	-2.6
Secondary school	41.6	42.9	-1.3
Post-secondary (technical or university)	29.2	25.4	3.8
Received a grant (%)	66.6	68.3	-1.7
Mean grant amount (for grantees only, JOD)	566	565	1
Baseline asset index			
Quartile 1 (%)	24.9	30.3	-5.4
Quartile 2 (%)	25.0	23.0	2.0
Quartile 3 (%)	25.1	23.9	1.2
Quartile 4 (%)	24.9	22.9	2.0
Baseline housing characteristics			
Owns home (%)	27.9	28.7	-0.8
Persons per room	1.9	1.9	0.0
Program site (%):			
Amman	26.2	25.0	1.2
Irbid	14.7	11.9	2.8
Kufrsoum	17.6	19.3	-1.7
Russeifa	15.0	15.9	-0.9
Zarqa	26.5	28	-1.5
Sample sizes	**1,235**	**757**	**n.a.**

Sources: NEF program activity data and NEF vulnerability assessment

Notes: n.a. = Not applicable. Youth are defined as being age 25 or under at the beginning of the program. The asset index is based on housing characteristics and durable goods ownership before Cohort 1 started the program, using a principal component analysis. We do not test for statistical significance between the analysis sample and population because those groups are not mutually exclusive.

Table A2b. Representativeness of the Cohort 1 grantee-only impact analysis sample

Characteristic	All Cohort 1 grantees	Cohort 1 grantee impact analysis sample	Difference
Demographic, household, and grant characteristics:			
Women (%)	83.4	90.6	-7.2
Youth (%)	26.5	18.6	7.9
Refugees (%)	28.3	30.3	-2
Mean age at the time of program intake (years)	35.6	37.6	-2.0
Head of household (%)	29.2	28.9	0.3
Has a disability (%)	7.3	9.0	-1.7
Literate (%)	97.4	96.7	0.7
Mean household size at time of program intake	5.3	5.3	0.0
Education level (%)			
Less than secondary education	23.8	25.6	-1.8
Secondary school	42.8	44.7	-1.9
Post-secondary (technical or university)	33.4	29.7	3.7
Mean grant amount (for grantees only, JOD)	566	565	1
Baseline asset index			
Quartile 1 (%)	23.8	27.7	-3.9
Quartile 2 (%)	24.5	20.1	4.4
Quartile 3 (%)	25	25.6	-0.6
Quartile 4 (%)	26.6	26.6	0
Baseline housing characteristics			
Owns home (%)	30.7	33.8	-3.1
Persons per room	1.8	1.8	0.0
Program site (%):			
Amman	27.2	25.8	1.4
Irbid	15.8	10.9	4.9
Kufrsoum	19.8	21.9	-2.1
Russeifa	12.9	14.1	-1.2
Zarqa	24.3	27.3	-3
Sample sizes	**823**	**488**	**n.a.**

Sources: NEF program activity data and NEF vulnerability assessment

Notes: n.a. = Not applicable. Youth are defined as being age 25 or under at the beginning of the program. The asset index is based on housing characteristics and durable goods ownership before Cohort 1 started the program, using a principal component analysis. We do not test for statistical significance between the analysis sample and population because those groups are not mutually exclusive.

Table A3a. Baseline equivalence of the treatment (Cohort 1) and comparison (Cohort 3) samples after matching

Characteristic	Cohort 1 impact analysis sample	Cohort 3 impact analysis sample	Difference
Demographic and household characteristics			
Women (%)	88.4	88.4	0.0
Youth (%)	23.5	23.5	0.0
Refugees (%)	35.0	35.0	0.0
Mean age at the time of Cohort 1 program intake (years)	36.8	35.1	1.7**
Head of household (%)	30.1	31.5	-1.4
Has a disability (%)	7.7	2.9	4.8**
Literate (%)	96.7	98.1	-1.4
Mean household size at time of Cohort 1 program intake	5.3	5.3	0.0
Education level (%)			
Less than secondary education	31.7	30.2	1.5
Secondary school	42.9	38.8	4.1
Post-secondary (technical or university)	25.4	31.1	-5.7*
Baseline asset index			
Quartile 1 (%)	30.3	30.3	0.0
Quartile 2 (%)	23.0	23.0	0.0
Quartile 3 (%)	23.9	23.9	0.0
Quartile 4 (%)	22.9	22.9	0.0
Baseline housing characteristics			
Owns home (%)	28.7	27.6	1.1
Persons per room	1.9	1.7	0.2**
Location at time of Cohort 1 program intake (%)			
Amman	25.0	25.0	0.0
Irbid	11.9	11.9	0.0
Kufrsoum	19.3	19.3	0.0
Russeifa	15.9	15.9	0.0
Zarqa	28.0	28.0	0.0
Sample sizes	**757**	**890**	**n.a.**

Sources: NEF program activity data and NEF vulnerability assessment

Notes: n.a. = Not applicable. Cohort 3 means and differences are estimated using coarsened exact matching weights. Youth are defined as being age 25 or under at the beginning of the program. The asset index is based on housing characteristics and durable goods ownership before Cohort 1 started the program, using a principal component analysis. The difference column may not exactly match the difference between group means due to rounding.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Table A3b. Baseline equivalence of the grantee-only treatment (Cohort 1) and comparison (Cohort 3) samples after matching

Characteristic	Cohort 1 impact analysis sample	Cohort 3 impact analysis sample	Difference
Demographic and household characteristics			
Women (%)	90.6	90.6	0.0
Youth (%)	18.6	18.6	0.0
Refugees (%)	30.3	30.3	0.0
Mean age at the time of Cohort 1 program intake (years)	37.6	35.5	2.1*
Head of household (%)	32.9	28.9	4
Has a disability (%)	9	3.1	5.9*
Literate (%)	98.5	96.7	1.8
Mean household size at time of Cohort 1 program intake	5.3	5.2	0.1
Education level (%)			
Less than secondary education	25.6	24	1.6
Secondary school	44.7	36.2	8.5**
Post-secondary (technical or university)	29.7	39.8	-10.1**
Baseline asset index /			
Quartile 1 (%)	27.7	27.7	0.0
Quartile 2 (%)	20.1	20.1	0.0
Quartile 3 (%)	25.6	25.6	0.0
Quartile 4 (%)	26.6	26.6	0.0
Baseline housing characteristics			
Owns home (%)	33.8	29.3	4.5
Persons per room	1.8	1.6	0.2**
Location at time of Cohort 1 program intake (%)			
Amman	25.8	25.8	0.0
Irbid	10.9	10.9	0.0
Kufrsoum	21.9	21.9	0.0
Russeifa	14.1	14.1	0.0
Zarqa	27.3	27.3	0.0
Sample sizes	**488**	**451**	**n.a.**

Sources: NEF program activity data and NEF vulnerability assessment

Notes: n.a. = Not applicable. Cohort 3 means and differences are estimated using coarsened exact matching weights. Youth are defined as being age 25 or under at the beginning of the program. The asset index is based on housing characteristics and durable goods ownership before Cohort 1 started the program, using a principal component analysis. The difference column may not exactly match the difference between group means due to rounding.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Table A3c. Baseline equivalence of the non-grantee-only treatment (Cohort 1) and comparison (Cohort 3) samples after matching

Characteristic	Cohort 1 impact analysis sample	Cohort 3 impact analysis sample	Difference
Demographic and household characteristics			
Women (%)	87.8	87.8	0.0
Youth (%)	28.8	28.8	0.0
Refugees (%)	36.6	36.6	0.0
Mean age at the time of Cohort 1 program intake (years)	35.9	34.6	1.3
Head of household (%)	29.8	30.7	-0.9
Has a disability (%)	2.9	5.4	-2.5
Literate (%)	98	98	0
Mean household size at time of Cohort 1 program intake	5.3	5.6	-0.3
Education level (%)			
Less than secondary education	36.5	39.5	-3
Secondary school	40.6	41.5	-0.9
Post-secondary (technical or university)	22.9	19	3.9
Baseline asset index			
Quartile 1 (%)	35.6	35.6	0
Quartile 2 (%)	24.9	24.9	0
Quartile 3 (%)	20.5	20.5	0
Quartile 4 (%)	19.0	19.0	0
Baseline housing characteristics			
Owns home (%)	25.5	23.9	1.6
Persons per room	2.1	1.9	0.2*
Location at time of Cohort 1 program intake (%)			
Amman	23.9	23.9	0.0
Irbid	14.6	14.6	0.0
Kufrsoum	14.1	14.1	0.0
Russeifa	19.5	19.5	0.0
Zarqa	27.8	27.8	0.0
Sample sizes	**205**	**283**	**n.a.**

Sources: NEF program activity data and NEF vulnerability assessment

Notes: n.a. = Not applicable. Cohort 3 means and differences are estimated using coarsened exact matching weights. Youth are defined as being age 25 or under at the beginning of the program. The asset index is based on housing characteristics and durable goods ownership before Cohort 1 started the program, using a principal component analysis. The difference column may not exactly match the difference between group means due to rounding.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Additional Tables: Section V

Table A4. Business practice scores for all cohorts, active businesses only

Sample	Sample size	Business records (0-7 points)	Financial planning (0-4 points)	Buying and stock management (0-3 points)	Marketing (0-7 points)	Total (0-21 points)	Total (%)
Full sample:							
All	**1,791**	**6.5**	**2.6**	**2.0**	**5.2**	**16.3**	**77.7**
By gender:							
Women	1,499	6.5	2.6	2.0	5.3	16.4	78.3
Men	292	6.2	2.5	1.9	5.1	15.7	74.9
By age:							
Youth	423	6.4	2.5	2.0	5.1	16.0	76.1
Adult	1,368	6.5	2.6	2.0	5.3	16.4	78.2
By refugee status:							
Refugees	534	6.5	2.5	1.9	5.3	16.3	77.4
Jordanians	1,257	6.5	2.6	2.1	5.2	16.3	77.8
By asset quartile:							
Quartile 1	440	6.5	2.7	2.0	5.4	16.6	78.9
Quartile 2	458	6.4	2.6	2.0	5.3	16.4	77.9
Quartile 3	452	6.5	2.5	2.0	5.2	16.2	77.4
Quartile 4	441	6.4	2.5	2.0	5.1	16.1	76.6

Source: IGA surveys

Notes: The sample is weighted to account for differences in cohort size. Scales are adapted from McKenzie and Woodruff (2021). Appendix B provides additional details on the practices included in each subscale. Youth are defined as being under age 25 at the beginning of the program. The asset index is based on housing characteristics and durable goods ownership from NEF's vulnerability assessment, using a principal component analysis for all three cohorts.

Table A5. Business metric for all cohorts, overall and by subgroup

Sample	Sample size	Active business: at least one sales transaction in the past 60 days (%)	Active business: no sales transaction but at least one purchase transaction in the past 60 days (%)	Active business: no sales or purchase transaction but productive activity in the past 60 days (%)	No active business but formally employed (%)	Total value of business metric (%)
All	1,838	96.2	0.8	0.3	0.2	97.6
By gender:						
Women	1,532	96.7	0.8	0.4	0.0	97.9
Men	306	93.7	1.3	0.0	1.2	96.1
By age:						
Youth	434	96.4	0.3	0.5	0.3	97.5
Adult	1,404	96.1	1.0	0.3	0.2	97.6
By refugee status:						
Refugees	550	96.0	0.5	0.3	0.2	97.1
Jordanians	1,288	96.3	1.0	0.3	0.2	97.8
By asset quartile:						
Quartile 1	455	95.4	0.6	0.6	0.0	96.6
Quartile 2	470	95.9	1.0	0.4	0.3	97.6
Quartile 3	463	96.4	1.0	0.2	0.3	97.9
Quartile 4	450	97.1	0.8	0.0	0.3	98.2

Source: IGA surveys

Notes: The sample is weighted to account for differences in cohort size. Subtotals may not match the total due to rounding. Youth are defined as being under age 25 at the beginning of the program. The asset index is based on housing characteristics and durable goods ownership from NEF's vulnerability assessment, using a principal component analysis for all three cohorts.

Table A6a. Business financial metrics for all cohorts, active businesses only (means)

Sample	Sample size	Number of days since last sale	Monthly revenues (JOD)	Monthly costs (JOD)	Estimated monthly profits (JOD)	Grantees with positive profits (%)	Monthly take-home business income (JOD)	Business savings (JOD)	Business debt (JOD)
All	1,791	8	343	210	133	88	98	226	75
By gender:									
Women	1,499	8	292	167	125	87	83	210	34
Men	292	7	603	430	172	90	175	305	283
By age:									
Youth	423	8	334	228	106	87	87	236	63
Adult	1,368	8	346	204	141	88	101	223	79
By refugee status:									
Refugees	534	8	329	184	144	93	104	158	76
Jordanians	1,257	8	349	221	128	86	95	255	75
By asset quartile:									
Quartile 1	440	8	326	188	137	90	104	174	49
Quartile 2	458	8	353	198	154	89	99	230	82
Quartile 3	452	8	370	242	127	86	102	251	76
Quartile 4	441	8	323	212	111	86	86	249	95

Source: IGA surveys

Note: The sample is weighted to account for differences in cohort size. Sample sizes vary slightly across outcomes. The sample sizes reported here are the number of grantees who reported any of the outcomes in the table. The incidence of missing values was low (1 percent or less). The asset index is based on housing characteristics and durable goods ownership from NEF's vulnerability assessment, using a principal component analysis for all three cohorts.

Table A6b. Business financial metrics for all cohorts, active businesses only (medians)

Sample	Sample size	Number of days since last sale	Monthly revenues (JOD)	Monthly costs (JOD)	Estimated monthly profits (JOD)	Monthly take-home income (JOD)	Business savings (JOD)	Business debt (JOD)
Full sample:								
All	1,791	3	206	95	100	55	100	0
By gender:								
Women	1,499	3	191	85	90	50	100	0
Men	292	2	400	150	171	130	120	0
By age:								
Youth	423	3	200	88	90	50	100	0
Adult	1,368	3	211	100	100	60	100	0
By refugee status:								
Refugees	534	3	202	95	107	70	80	0
Jordanians	1,257	3	208	96	94	50	150	0
By asset quartile:								
Quartile 1	440	3	200	95	102	70	85	0
Quartile 2	458	3	210	100	100	60	100	0
Quartile 3	452	3	220	100	97	60	120	0
Quartile 4	441	3	203	90	90	50	150	0

Source: IGA surveys

Note: The sample is weighted to account for differences in cohort size. Sample sizes vary slightly across outcomes. The sample sizes reported here are the number of grantees who reported any of the outcomes in the table. The incidence of missing values was low (1 percent or less). The asset index is based on housing characteristics and durable goods ownership from NEF's vulnerability assessment, using a principal component analysis for all three cohorts.

Table A7. Key IGA outcomes by cohort

Sample	Cohort 1	Cohort 2	Cohort 3	All Cohorts
DIB business metric				
Overall business metric	98.5	96.0	98.4	97.6
95 percent confidence interval for the business metric	[97.8, 99.2]	[95.2, 96.8]	[97.9, 98.8]	[97.2, 97.9]
Active business: at least one sales transaction in the past 60 days (%)	97.2	94.7	96.9	96.2
Active business: no sales transaction but at least one purchase transaction in the past 60 days (%)	1.2	0.5	1.0	0.8
Active business: no sales or purchase transaction but productive activity in past 60 days (%)	0.2	0.8	0.0	0.3
No active business but formally employed (%)	0.0	0.0	0.5	0.2
Sample size:	**603**	**626**	**609**	**1,838**
Business financial metrics (active businesses only)				
Mean number of days since last sale	9	8	7	8
Mean monthly revenue (JOD)	356	314	359	343
Mean monthly costs (JOD)	225	189	218	210
Mean monthly profits (JOD)	130	126	140	133
Mean monthly take-home business income (JOD)	89	96	104	98
Mean business savings (JOD)	290	186	222	226
Mean business debt (JOD)	92	84	58	75
Sample size:	**594**	**601**	**599**	**1,794**
Business practice scores (active businesses only)				
Business records (0-7)	6.5	6.6	6.3	6.5
Financial planning (0-4)	2.7	2.7	2.4	2.6
Buying and stock management (0-3)	2.1	2.1	1.9	2.0
Marketing (0-7)	5.2	5.5	5.1	5.3
Total (0-21)	16.4	16.9	15.7	16.3
Total (%)	78.3	80.5	75.0	77.7
Sample size:	**594**	**601**	**599**	**1,794**

Source: Cohort 1 and 2 IGA surveys

Additional Tables: Section VI

Table A8. Business metric two years post-grant among Cohort 1 grantees, by subgroup

Sample	Sample size	Active business: at least one sales transaction in the past 60 days (%)	Active business: no sales transaction but at least one purchase transaction in the past 60 days (%)	Active business: no sales or purchase transaction but productive activity in the past 60 days (%)	No active business but formally employed (%)	Total value of medium-term version of business metric (%)
Full sample:						
All	**550**	**74.7**	**1.1**	**0.0**	**0.5**	**76.4**
By gender:						
Women	481	75.9	1.0	0.0	0.4	77.3
Men	69	66.7	1.4	0.0	1.4	69.6
By age:						
Youth	123	72.4	0.0	0.0	0.8	73.2
Adult	427	75.4	1.4	0.0	0.5	77.3
By refugee status:						
Refugees	174	73.0	0.0	0.0	0.0	73.0
Jordanians	376	75.5	1.6	0.0	0.8	77.9
By asset quartiles						
Quartile 1	150	75.3	0.7	0.0	0.0	76.0
Quartile 2	118	69.5	1.7	0.0	0.0	71.2
Quartile 3	138	72.5	0.7	0.0	1.4	74.6
Quartile 4	144	80.6	1.4	0.0	0.7	82.6

Source: Impact survey

Notes: Subtotals may not match the total due to rounding. Youth are defined as being under age 25 at the beginning of the program.

Table A9a. Business financial metrics for Cohort 1 grantees with active businesses two years post-grant (means)

Sample	Sample size	Number of days since last sale	Monthly revenues (JOD)	Monthly costs (JOD)	Estimated monthly profits (JOD)	Grantees with positive profits (%)	Current monthly take-home income (JOD)	Business savings (JOD)	Business debt (JOD)
Full sample:									
All	**417**	**11**	**251**	**175**	**81**	**80.1**	**91**	**235**	**111**
By gender:									
Women	370	11	222	154	75	78.9	82	232	47
Men	47	11	473	340	133	89.4	162	257	610
By age:									
Youth	89	11	287	158	136	83.7	81	255	192
Adult	328	10	241	180	67	79.2	94	230	89
By refugee status:									
Jordanian	290	10	271	197	79	79.3	91	264	112
Refugees	127	12	207	123	87	82.1	92	170	111
By asset quartiles									
Quartile 1	114	11	202	123	82	82.7	87	166	41
Quartile 2	84	11	244	153	90	78.3	87	194	209
Quartile 3	101	10	289	250	41	75.5	96	238	116
Quartile 4	118	10	272	176	111	83.0	95	330	104

Source: Impact survey

Note: Sample sizes vary slightly across outcomes. The sample sizes reported here are the number of grantees who reported any of the outcomes in the table. The incidence of missing values was low (1 percent or less).

Table A9b. Business financial metrics for Cohort 1 grantees with active businesses two years post-grant (medians)

Sample	Sample size	Number of days since last sale	Monthly revenues (JOD)	Monthly costs (JOD)	Estimated monthly profits (JOD)	Grantees with positive profits (%)	Current monthly take-home income (JOD)	Business savings (JOD)	Business debt (JOD)
Full sample:									
All	**417**	**4**	**150**	**66**	**60**	**80.1**	**50**	**120**	**0**
By gender:									
Women	370	4	150	60	50	78.9	50	120	0
Men	47	4	300	100	150	89.4	150	115	0
By age:									
Youth	89	6	150	50	70	83.7	50	100	0
Adult	328	4	150	70	60	79.2	50	120	0
By refugee status:									
Jordanian	290	3	150	70	70	79.3	50	150	0
Refugees	127	7	150	50	50	82.1	50	70	0
By asset quartiles									
Quartile 1	114	5	150	50	50	82.7	50	80	0
Quartile 2	84	4.5	120	82.5	50	78.3	50	100	0
Quartile 3	101	4.5	160	80	80	75.5	50	150	0
Quartile 4	118	4	147	50	70	83.0	50	300	0

Source: Impact survey

Notes: Sample sizes vary slightly across outcomes. The sample sizes reported here are the number of grantees who reported any of the outcomes in the table. The incidence of missing values was low (1 percent or less).

Table A10. Impacts on business ownership and annual household income

Outcome	Treatment mean	Comparison mean	Difference	Effect size (SD)
Business ownership				
Owns a business (%)	63	35	28**	0.59**
Average number of businesses owned, among households that own a business	1.2	1.0	0.2**	1.46**
Annual income, overall and by source				
Total household income (JOD)	4,681	4,007	674**	0.24**
Business income (JOD)	841	343	498**	0.58**
Income from wages (JOD)	2,583	2,489	94	0.04
Income from pensions	608	562	46	0.03
Social assistance income (JOD)	495	395	100*	0.14*
Remittances, family support, and income from assets (JOD)	154	219	-65	-0.07
Sample size	**757**	**890**	**n.a.**	**n.a.**

Source: Impact survey

Notes: Samples include grantees and non-grantees. Comparison means and treatment-comparison differences are estimated using coarsened exact matching weights and regression adjusted with controls for household size and its square, age at the time of program entry for Cohort 1 and its square, the continuous asset index, and education level (less than secondary, secondary, or more than secondary). We conducted top- and bottom-coding to the 99th percentile separately by cohort for each income category to avoid outliers unduly influencing the findings. Effect sizes were calculated using the standard deviation of the comparison group.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Table A11. Impacts on total annual household consumption for the full sample, overall and by consumption category

Sample	Treatment mean (JOD)	Comparison mean (JOD)	Difference (JOD)	Effect size (SD)
Annual consumption				
Consumption metric: Total household consumption	6,983	6,347	636**	0.22**
Direct consumption (proxy for expenditure)[a]	6,053	5,498	555**	0.23**
Per-capita household consumption[b]	1,936	1,819	116**	0.13**
Annual consumption, by category				
Food items	2,255	1,996	259**	0.27**
Non-food items	2,800	2,474	326**	0.18**
Durables	320	269	51*	0.11*
Housing	1,608	1,609	-1	0
Sample size	**757**	**890**	**n.a.**	**n.a.**

Source: Impact survey

Notes: n.a. = Not applicable. Samples include grantees and non-grantees. Comparison means and treatment-comparison differences are estimated using coarsened exact matching weights and regression adjusted with controls for household size and its square, age at the time of program entry for Cohort 1 and its square, the continuous asset index, and education level (less than secondary, secondary, or more than secondary). We conducted top- and bottom-coding separately by cohort for each consumption category to avoid outliers unduly influencing the findings. Food, nonfood, and housing were top and bottom coded to the 99th and 1st percentiles, respectively, by cohort. Durables goods were top coded to the 99th percentile by cohort. The difference column may not exactly match the difference between group means due to rounding. Effect sizes were calculated using the standard deviation of the comparison group.

[a] Direct consumption includes the value of food consumption, expenditure on non-food goods and services, and rent payments; it excludes durable goods and estimated rent for owner-occupied housing, which are included in the primary consumption measure.

[b] Adult equivalent calculations use the OECD equivalence scale which assigns a value of 1 to the first household member, 0.7 for each additional adult aged 14 or older, and 0.5 to each child.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Table A12. Impacts on food security and livelihoods coping strategies

Sample	Treatment mean	Comparison mean	Difference	Effect size (SD)
Food security: reduced coping strategies index (rCSI)				
Reduced Coping Strategy Index (0-56)	12.1	13.5	-1.4*	-0.12*
Sub-categories (%)				
No coping strategies	9.0	5.7	3.3*	0.14*
Low severity strategies only	38.8	34.0	4.9	0.10
Moderate severity strategies	14.5	18.5	-4	-0.10
Severe coping strategies	37.7	41.9	-4.2	-0.08
Sample size	**677**	**827**	**n.a.**	**n.a.**
Livelihoods coping strategies index				
The livelihoods coping strategies index (0-24)	5.2	5.8	-0.6**	-0.15**
Sub-categories (%)				
No coping strategies	25.4	17.7	7.7**	0.2**
Stress-level strategies	0.7	0.6	0.2	0.02
Crisis-level strategies	55.8	63.9	-8.1**	-0.17**
Emergency coping strategies	18.0	17.8	0.2	0
Sample size	**677**	**845**	**n.a.**	**n.a.**

Source: Impact survey

Notes: n.a. = Not applicable. Samples include grantees and non-grantees. Comparison means and treatment-comparison differences are estimated using coarsened exact matching weights and regression adjusted with controls for household size and its square, age at the time of program entry for Cohort 1 and its square, the continuous asset index, and education level (less than secondary, secondary, or more than secondary). We imputed "don't know" and "refuse" responses with the mean by cohort and refugee status for up to 20 percent of the items in the scale, and treated a respondent as missing if they declined to respond to more than 20 percent of items. Nonresponse rates were similar across cohorts. The difference column may not exactly match the difference between group means due to rounding. Effect sizes were calculated using the standard deviation of the comparison group.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Table A13. Impacts on total annual household income, by subgroup

Sample	Treatment sample size	Comparison sample size	Treatment mean (JOD)	Comparison mean (JOD)	Difference (JOD)	Effect size (SD)
By gender						
Women	669	765	4,499	3,969	529**	0.19**
Men	88	125	6,064	4,314	1,750**	0.51**
By age						
Youth	178	281	5,217	4,058	1,160**	0.42**
Adult	579	609	4,516	4,019	497**	0.17**
By refugee status						
Refugees	265	288	3,650	2,947	703**	0.38**
Jordanians	492	602	5,235	4,585	650**	0.21**
By asset quartile						
Quartile 1	229	301	3,674	2,914	760**	0.36**
Quartile 2	174	229	4,106	3,682	425	0.17
Quartile 3	181	179	5,022	4,296	726*	0.29*
Quartile 4	173	181	6,233	5,349	884*	0.25*

Source: Impact survey

Notes: Samples include grantees and non-grantees. Comparison means and treatment-comparison differences are estimated using coarsened exact matching weights and regression adjusted with controls for household size and its square, age at the time of program entry for Cohort 1 and its square, the continuous asset index, and education level (less than secondary, secondary, or more than secondary). We conducted top- and bottom-coding to the 99th percentile separately by cohort for each income category to avoid outliers unduly influencing the findings. The difference column may not exactly match the difference between group means due to rounding. Effect sizes were calculated using the standard deviation of the comparison group. Differences in impacts across related subgroups were not statistically significant except for the difference between youth and adults, which was statistically significant at the .05 level, and between men and women, which was statistically significant at the 0.01 level.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Table A14. Impacts on annual household income, by income source and subgroup

Sample	Total impact (JOD)	Disaggregated impact, by source (JOD)				
		Business income	Wages	Pensions	Social assistance income	Other income
All	674**	498**	94	46	100*	-65
By gender						
Women	529**	514**	-24	-12	98*	-47
Men	1,750**	386	1,000*	422*	112	-170
By age						
Youth	1,160**	399**	250	445**	16	49
Adult	497**	535**	14	-82	121*	-92
By refugee status						
Refugees	703**	364**	148	-5	113	83
Jordanians	650**	569**	96	48	85	-148*
By asset quartile						
Quartile 1	760**	358**	120	69	175*	39
Quartile 2	425	296*	-36	-38	69	134
Quartile 3	726*	610**	249	-79	93	-149
Quartile 4	884*	782**	81	286	43	-308*

Source: Impact survey

Notes: Samples include grantees and non-grantees. Business income includes take-home business income for all household businesses. Wages includes wage income for all household members, including from both informal and formal employment. Other income includes income from assets, remittances, and support from family and neighbors. Comparison means and treatment-comparison differences are estimated using coarsened exact matching weights and regression adjusted with controls for household size and its square, age at the time of program entry for Cohort 1 and its square, the continuous asset index, and education level (less than secondary, secondary, or more than secondary). We conducted top- and bottom-coding to the 99th percentile separately by cohort for each income category to avoid outliers unduly influencing the findings. The difference column may not exactly match the difference between group means due to rounding.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Table A15. Impacts on total annual household consumption, by subgroup

Sample	Treatment sample size	Comparison sample size	Treatment mean (JOD)	Comparison mean (JOD)	Difference (JOD)	Effect size (SD)
By gender						
Women	669	765	6,908	6,268	640**	0.23**
Men	88	125	7,555	6,976	578	0.18
By age						
Youth	178	281	7,410	6,433	976**	0.33**
Adult	579	609	6,852	6,310	542**	0.19**
By refugee status						
Refugees	265	288	5,173	4,959	214	0.11
Jordanians	492	602	7,958	7,100	858**	0.29**
By asset quartile						
Quartile 1	229	301	5,434	5,001	433*	0.21*
Quartile 2	174	229	6,061	5,914	147	0.06
Quartile 3	181	179	7,493	6,650	844**	0.31**
Quartile 4	173	181	9,427	8,252	1,174**	0.37**

Source: Impact survey

Notes: n.a. = Not applicable. Comparison means and treatment-comparison differences are estimated using coarsened exact matching weights and regression adjusted with controls for household size and its square, age at the time of program entry for Cohort 1 and its square, the continuous asset index, and education level (less than secondary, secondary, or more than secondary). We conducted top- and bottom-coding separately by cohort for each consumption category to avoid outliers unduly influencing the findings. Food, nonfood, and housing were top and bottom coded to the 99th and 1st percentiles, respectively, by cohort. Durables goods were top coded to the 99th percentile by cohort. Differences in impacts across related subgroups were not statistically significant except for the differences across asset quartiles and between Jordanians and refugees, which were both statistically significant at the 0.05 level. The difference column may not exactly match the difference between group means due to rounding. Effect sizes were calculated using the standard deviation of the comparison group.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Table A16. Impacts on the reduced coping strategies index, by subgroup

Sample	Treatment sample size	Comparison sample size	Treatment mean	Comparison mean	Difference	Effect size (SD)
By gender						
Women	602	709	12.2	13.5	-1.3	-0.11
Men	75	118	10.6	12.7	-2.1	-0.19
By age						
Youth	160	267	9.4	12.5	-3.2**	-0.31**
Adult	517	560	12.9	13.6	-0.7	-0.06
By refugee status						
Refugees	228	261	16.9	16.2	0.8	0.07
Jordanians	449	566	9.6	12.2	-2.6**	-0.24**
By asset quartile						
Quartile 1	201	272	16.1	17.2	-1.1	-0.09
Quartile 2	151	217	14.2	13.2	1	0.10
Quartile 3	163	165	9.9	12.7	-2.8*	-0.26*
Quartile 4	162	173	7.3	10.8	-3.5**	-0.34**

Source: impact survey

Notes: Samples include grantees and non-grantees. Comparison means and treatment-comparison differences are estimated using coarsened exact matching weights and regression adjusted with controls for household size and its square, age at the time of program entry for Cohort 1 and its square, the continuous asset index, and education level (less than secondary, secondary, or more than secondary). We imputed "don't know" and "refuse" responses with the mean by cohort and refugee status for up to 20 percent of the items in the scale, and treated a respondent as missing if they declined to respond to more than 20 percent of items. Nonresponse rates were similar across cohorts. The difference column may not exactly match the difference between group means due to rounding. Differences in impacts across related subgroups were not statistically significant except for the differences across asset quartiles, which were statistically significant at the .05 level, and between Jordanians and refugees, which was statistically significant at the 0.01 level. Effect sizes were calculated using the standard deviation of the comparison group.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Table A17. Impacts on the livelihoods coping index, by subgroup

Sample	Treatment sample size	Comparison sample size	Treatment mean	Comparison mean	Difference	Effect size (SD)
By gender						
Women	602	730	5.2	-0.6**	0.12	602
Men	75	115	5.0	-0.3	0.03	75
By age						
Youth	160	268	4.6	-1.3**	0.16	160
Adult	517	577	5.4	-0.4	0.10	517
By refugee status						
Refugees	229	266	7.6	0.0	0.14	229
Jordanians	448	579	3.9	-1.0**	0.13	448
By asset quartile						
Quartile 1	203	280	7.1	0.0	0.15	203
Quartile 2	151	219	5.6	-0.7	0.02	151
Quartile 3	159	172	4.4	-0.7	0.19	159
Quartile 4	164	174	3.2	-1.6**	0.24	164

Source: Impact survey

Notes: Samples include grantees and non-grantees. Comparison means and treatment-comparison differences are estimated using coarsened exact matching weights and regression adjusted with controls for household size and its square, age at the time of program entry for Cohort 1 and its square, the continuous asset index, and education level (less than secondary, secondary, or more than secondary). We imputed "don't know" and "refuse" responses with the mean by cohort and refugee status for up to 20 percent of the items in the scale, and treated a respondent as missing if they did not respond to more than 20 percent of items. The difference column may not exactly match the difference between group means due to rounding. Differences in impacts across related subgroups were not statistically significant except for the difference between Jordanians and refugees, which was statistically significant at the 0.05 level. Effect sizes were calculated using the standard deviation of the comparison group.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Table A18. Impacts on the baseline asset index, overall and by subgroup, in terms of baseline standard deviations

Sample	Treatment sample size	Comparison sample size	Treatment mean (SD)	Comparison mean (SD)	Difference (SD)
All	757	889	0.43	0.30	0.13**
By gender					
Women	669	764	0.45	0.31	0.14**
Men	88	125	0.29	0.26	0.04
By age					
Youth	178	280	0.46	0.26	0.2*
Adult	579	609	0.42	0.31	0.11*
By refugee status					
Refugees	265	288	-0.47	-0.56	0.09
Jordanians	492	601	0.92	0.77	0.15*
By asset quartile					
Quartile 1	229	300	-0.43	-0.55	0.12
Quartile 2	174	229	0.07	0.05	0.02
Quartile 3	181	179	0.80	0.61	0.19
Quartile 4	173	181	1.55	1.33	0.22*

Source: Impact survey

Notes: Samples include grantees and non-grantees. Comparison means and treatment-comparison differences are estimated using coarsened exact matching weights and regression adjusted with controls for household size and its square, age at the time of program entry for Cohort 1 and its square, the continuous asset index, and education level (less than secondary, secondary, or more than secondary). The difference column may not exactly match the difference between group means due to rounding. Differences in impacts across related subgroups were not statistically significant. Effect sizes were calculated using the standard deviation of the comparison group.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Table A19. Impacts on subjective well-being and women's empowerment

Outcome	Treatment sample size	Comparison sample size	Treatment mean	Comparison mean	Difference	Effect size (SD)
Subjective well-being						
Satisfaction with life scale (maximum 35)	625	772	19.2	19.1	0.2	0.04
Self-confidence scale (maximum 30)	662	834	21.7	21.9	-0.2	-0.07
Percent of women with a moderate or high influence over decisions regarding… (%)						
Use of household business resources (households with businesses only)	407	260	84.8	86.8	-2.1	-0.06
Large household purchases	605	736	78.3	73.6	4.7	0.11
Their own health care	606	736	83.8	82.5	1.4	0.04
Visits to family or relatives	607	738	76.1	72.7	3.4	0.08
Freedom of movement scale						
Women's freedom of movement (maximum 4)	584	697	2.6	2.6	0.0	0.0

Source: Impact survey

Notes: Samples include grantees and non-grantees. Comparison means and treatment-comparison differences are estimated using coarsened exact matching weights and regression adjusted with controls for household size and its square, age at the time of program entry for Cohort 1 and its square, the continuous asset index, and education level (less than secondary, secondary, or more than secondary. The difference column may not exactly match the difference between group means due to rounding. Effect sizes were calculated using the standard deviation of the comparison group. No treatment-comparison difference are significantly different from zero at the .05 level or better.

Table A20. Impacts on education outcomes, among households with members in the relevant age range

Outcome	Treatment sample size	Comparison sample size	Treatment mean	Comparison mean	Difference	Effect size (SD)
Compulsory education (Ages 6-16)						
All enrolled	510	582	92.7	95.7	-3.0*	-0.15*
Full attendance	389	402	75.3	71.5	3.8	0.09
Non-compulsory secondary education (Ages 17-18)						
All enrolled	262	264	85.7	76.9	8.7*	0.22*
Full attendance	166	143	68.7	62.7	6	0.12
Postsecondary education (Ages 19-24)						
Any enrolled	392	448	27.6	27.9	-0.3	-0.01

Source: Impact survey

Notes: Samples include grantees and non-grantees. Comparison means and treatment-comparison differences are estimated using coarsened exact matching weights and regression adjusted with controls for household size and its square, age at the time of program entry for Cohort 1 and its square, the continuous asset index, and education level (less than secondary, secondary, or more than secondary Attendance sample sizes are substantially lower because they exclude households where no eligible students were enrolled and those who were surveyed during school holidays. The difference column may not exactly match the difference between group means due to rounding. Effect sizes were calculated using the standard deviation of the comparison group.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Table A21. Impacts on measures of economic well-being for grantees only

Outcome	Treatment mean	Comparison mean	Difference	Effect size (SD)
Annual household income (JOD)				
Total household income	**5,158**	**4,275**	**883****	**0.32****
Household business income	1,152	410	742**	0.86**
Annual household consumption (JOD)				
Total household consumption	**7,401**	**6,456**	**945****	**0.36****
Total household consumption per adult equivalent	2,059	1,843	216**	0.28**
Food consumption	2,365	2,014	351**	0.37**
Nonfood consumption	3,006	2,519	488**	0.30**
Durable goods	369	305	64*	0.15*
Housing	1,661	1,619	42	0.07
Debt, savings, and household assets				
Household savings (JOD)	132	63	70*	0.21*
Household debt (JOD)	3,295	3,724	-429	-0.07
Household asset index terms of baseline standard deviations	0.66	0.37	0.30**	n.a.
Sample size for all preceding variables	**488**	**451**	**n.a.**	**n.a.**
Household and coping strategies				
Reduced coping strategy index (0-56)	11.2	12.0	-0.8	-0.08
Livelihood coping strategies index (0-19)	4.7	5.2	-0.5	-0.12
Sample size for coping strategies	**432**	**429**	**n.a.**	**n.a.**

Source: Impact survey

Notes: n.a. = not applicable. Comparison means and treatment-comparison differences are estimated using coarsened exact matching weights and regression adjusted with controls for household size and its square, age at the time of program entry for Cohort 1 and its square, the continuous asset index, and education level (less than secondary, secondary, or more than secondary. The difference column may not exactly match the difference between group means due to rounding. Effect sizes were calculated using the standard deviation of the comparison group.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Table A22. Impacts on measures of economic well-being for non-grantees only

Outcome	Treatment mean	Comparison mean	Difference	Effect size (SD)
Annual household income (JOD)				
Total household income	**3,588**	**3,647**	**-60**	**-0.02**
Household business income	132	329	-197*	-0.19*
Annual household consumption (JOD)				
Total household consumption	**6,110**	**5,993**	**116**	**0.04**
Total household consumption per adult equivalent	1,696	1,686	10	0.01
Food consumption	1,989	1,932	57	0.06
Nonfood consumption	2,376	2,270	107	0.06
Durable goods	234	240	-6	-0.01
Housing	1,509	1,551	-42	-0.08
Debt, savings, and household assets				
Household savings (JOD)	180	40	140	0.88
Household debt (JOD)	2,504	3,507	-1003*	-0.17*
Household asset index in terms of baseline standard deviations	0.04	0.13	-0.1	-0.07
Sample size for all preceding variables	**205**	**283**	**n.a.**	**n.a.**
Household and coping strategies				
Reduced coping strategy index (0-56)	14.6	14.5	0.1	0.01
Livelihood coping strategies index (0-19)	6.2	6.5	-0.3	-0.07
Sample size	**177**	**264**	**n.a.**	**n.a.**

Source: Impact survey

Notes: n.a. = not applicable. Comparison means and treatment-comparison differences are estimated using coarsened exact matching weights and regression adjusted with controls for household size and its square, age at the time of program entry for Cohort 1 and its square, the continuous asset index, and education level (less than secondary, secondary, or more than secondary. The difference column may not exactly match the difference between group means due to rounding. Effect sizes were calculated using the standard deviation of the comparison group.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Appendix B: Technical details for the IGA validation study

A. Sampling

To conduct the IGA survey sampling for each cohort, we first divided grantees into strata that were defined by mutually exclusive combinations of program site (five different locations), gender (men and women), age group (youth and adults), and refugee status (Jordanians and refugees). We then selected a random sample of 600 grantees the cohort, distributed across strata in proportion to strata's share of the population of grantees. We designated these 600 grantees as the primary sample and the remaining grantees as potential replacements.

When respondents from the primary sample were unreachable, unwilling to participate, or unavailable, we selected replacements from a randomly sorted list of potential replacements from the same stratum. If replacements from that stratum were exhausted, we drew replacements from the most similar stratum with replacements remaining. The average response rate for the primary sample across cohorts was 84 percent. Drawing on replacements, we ultimately achieved a sample size of 603 Cohort 1, 624 Cohort 2, and 607 Cohort 3 respondents, for a total of 1,834 respondents.[21] For Cohorts 2 and 3, the business metric also includes two additional grantees per cohort who declined to participate in the survey because they reported their business had closed.

B. Data collection approach

The data collection team participated in a three-day in-person training on the survey prior to commencing data collection in each cohort. The training included training on the protection of human subjects and vulnerable populations, a detailed review of the survey questions, and practice conducting interviews using a series of respondent scenarios. About a dozen different enumerators conducted each round of the survey in teams of two or three, each with the support of a logistical coordinator who coordinated with potential respondents while in the field. The study used a verbal informed consent statement and procedure approved by the King Hussein Cancer Center Institutional Review Board (IRB) in Jordan and Health Media Lab IRB in the United States. **Table B.1** summarizes details of the IGA survey.

[21] The sample size for Cohort 2 substantively exceeded the original target of 600 as additional follow-up attempts in the final days of data collection with initial nonrespondents proved successful.

Table B.1. Details on IGA survey data collection

Cohort	Cohort 1	Cohort 2	Cohort 3	Overall
Data collection dates	May 24, 2023 – June 22, 2023	May 8, 2024 - June 13, 2023	May 26, 2025 – June 24, 2025	n.a.
Number of enumerators	11	11	9	n.a.
Mean survey duration	15 minutes	14 minutes	17 minutes	15 minutes
Mean follow-up period from grant receipt (minimum – maximum)	10.3 months (9.9 – 10.8 months)	11.2 months (10.8 – 12.1 months)	8.3 months (7.8 – 9.0 months)	9.8 months n.a.
Sample size	603 grantees	626 grantees	609 grantees	1,838 grantees
Response rate	85 percent	86 percent	80 percent	84 percent

C. IGA metric calculation

The business metric is a count of the number of grantees with active IGAs divided by the sample size. For calculations of the confidence intervals around the business metric, the analysis accounts for the proportion of grantees in each sampling stratum who responded to the survey by incorporating a finite population correction (FPC). The FPC narrows the confidence intervals as the proportion of grantees included in the survey sample approaches 100 percent. Other statistics provided in this report were calculated using simple descriptive methods.

The aggregate business metric incorporates cohort-specific FPCs and cohort-level weights that account for the size of each cohort relative to the total cumulative population of grantees. Thus, the estimated business metric is representative of the full population of grantees.

D. Business practice scores

We assessed grantees' self-reported business management practices by scoring them in four separate business practice domains (McKenzie and Woodruff 2016); as mentioned earlier, these scores have been found to be positively correlated with business sales and profits across several low- and middle-income countries. Below we describe the calculation of the scores in each domain, which we summed to generate the total business practices score.

We also provide Cronbach's alpha for each score, which is a measure of scale reliability across different items, ranging from 0 to 1. An alpha of 1 corresponds to perfect alignment across items (for example, on a two-item scale, one item always equals 1 when another equals 0, and vice versa). The alpha for the total score is 0.52.

Business records[22]

The business records score adds one point for each of the following, if the respondent:

1. Keeps written business records;
2. Records every sale and purchase made by the business;
3. Can use business records to see how much cash the business has on hand;
4. Uses their records regularly to know whether sales of a particular good are increasing or decreasing;
5. Has worked out the cost to the business of the most common good or service they sell;
6. Knows which good or service they make the most profit per unit from selling; and
7. Has a written monthly budget for the business.

Maximum score: 7 / **Cronbach's alpha:** 0.38

Marketing

The marketing score adds one point for each of the following that the respondent reports doing.

In the last three months:

8. Visited at least one of their competitors' businesses to see what prices they are charging;
9. Visited at least one of their competitors' businesses to see what goods or services they are offering;
10. Asked their existing customers whether there are any other goods or services they would like the respondent to sell;
11. Talked with a former customer to find out why they have stopped buying from their business;
12. Asked a supplier about which products are selling well in their industry;
13. Used a special offer to attract customers; and

In the last six months:

14. Did any form of advertising.

Maximum score: 7 / **Cronbach's Alpha:** 0.61

Buying and stock management

The buying and stock management score adds one point for each of the following, if the respondent:

1. Attempted to negotiate with a supplier for a lower price for goods or raw material within the last three months;
2. Compared prices or quality from alternate suppliers against current suppliers within the last three months; and
3. Runs out of supplies or raw materials for their business less than once a month.

[22] We modified the original scale to remove one item: *"If you wanted to apply for a bank loan, and were asked to provide records to show that you have enough money left each month after paying business expenses to repay a loan, would your records allow you to document this to the bank?"* Although the question is designed to be hypothetical, refugee respondents struggled to respond to this item during pretesting and piloting because of a perceived lack of access to formal financial systems.

Maximum score: 3 / **Cronbach's Alpha:** 0.31

Financial planning

The financial planning score adds one point for each of the following, if the respondent:

1. Reviews the financial performance of their business and analyzes where there are areas for improvement at least once per month;
2. Has a specific, quantitative target for sales over the next year;
3. Compares actual to targeted sales at least once per month; and
4. Has a written budget of the likely costs their business will have to face over the next year.

Maximum score: 4 / **Cronbach's Alpha:** 0.51

Appendix C: Technical details for the impact evaluation

A. Impact evaluation design

The impact evaluation leverages the cohort-based rollout of the treatment to compare the outcomes of participants in Cohort 1 about 23 months after grant disbursement with the situation of a matched sample of those who recently started the program as part of Cohort 3. This matched comparison design aims to limit the possibility that differences in outcomes between the two cohorts are driven by differences in participant characteristics, making it more plausible to attribute these differences to the impact of the program. Specifically, because Cohorts 1 and 3 were selected for the program using a similar approach, we expect them to be broadly similar in terms of vulnerability and unobserved characteristics related to program participation and outcomes (for example, motivation and entrepreneurial spirit).

To further limit the possibility that differences in outcomes between the two cohorts are driven by differences in observed demographic and economic participant characteristics, we implemented coarsened exact matching (CEM) (Iacus et al. 2012). This method divides members of the two cohorts into mutually exclusive groups, defined by a combination of participant demographic characteristics and self-reported household assets and characteristics in late 2021 when Cohort 1 was selected.[23] For example, one possible stratum is female Jordanians who were under 25, received the program at the Irbid *Siraj* center, were in the lowest quartile of the sample in terms of household assets in late 2021, and had a household size of 3 or fewer members at the time. We then reweighted the comparison group (Cohort 3) so that its distribution across strata is identical to that of the treatment group (Cohort 1), making it as similar as possible to the treatment group in terms of the matching characteristics. Thus, all comparison group observations in the same stratum receive the same weight. In this way, the comparison group is adjusted so that it is as similar as possible to the treatment group in terms of the matching characteristics, enabling us to produce credible impact estimates.

B. Matching and sampling approach

The CEM approach is applied to a sample of survey respondents with data on the relevant characteristics and outcomes. It typically results in some sample loss because respondents in CEM strata that have only treatment or control observations (that is, unmatched individuals) are dropped. In this study, we planned to survey a sample of participants rather than all participants given our sample size targets and wanted to minimize subsequent sample loss due to unmatched respondents at the analysis stage. Therefore, we conducted an initial stage of matching before sampling that was intended to focus the sample on individuals who were more likely to be matched, while recognizing that the final matching could only be conducted once we knew who in the sample responded to the survey.

Specifically, we applied the following steps to select the survey sample:

[23] We obtained most of the information about matching characteristics from NEF's vulnerability assessment, which was used to screen participants shortly before they were selected (between February and April 2022 for Cohort 1 and between January and February 2024 for Cohort 3). For characteristics that were not time invariant, such as household assets, NEF asked these questions retrospectively for Cohort 3 about their situation in late 2021.

1. We conducted initial CEM using the following matching characteristics: gender, refugee status, youth status (25 years or younger, or older than 25), program site, household size (3 or less, 5–8 or 9 or more) education level (less then secondary, secondary, or more than secondary), and asset index quartile. We selected these matching variables to optimize baseline balance while retaining as much of the sample as possible. To estimate the asset index, we conducted a principal components analysis based on pre-program housing characteristics and durable goods owned at the time when Cohort 1 was starting the program.[24] Out of all program participants in the two cohorts, we matched 1,189 of the 1,235 participants in Cohort 1 (96 percent) and 2,231 of the 2,472 participants in Cohort 3 (90 percent) using this approach. [25]
2. We included all 1,189 matched Cohort 1 participants (1,189) in our survey sample because the cohort was smaller than Cohort 3 and because we anticipated higher levels of nonresponse given that these participants had left the program almost two years prior to the survey.
3. We then randomly selected a sample of 929 out of the 2,231 matched Cohort 3 participants as our primary sample, allocating this sample across strata defined by mutually exclusive combinations of gender, refugee status, youth status, and program site, in proportion to strata's share of the population of participants.[26] The rest of the matched Cohort 3 participants (1,302) served as replacements in the case of nonresponse. Specifically, when participants from the primary sample were unreachable, unwilling to participate, or unavailable, we selected replacements from a randomly sorted list of potential replacements from the same stratum. If replacements from that stratum were exhausted, we drew replacements from the most similar stratum with replacements remaining.

C. Data collection approach

1. Survey development

Our survey development approach sought to ensure that the impact survey captured as much of respondents' household consumption as possible without become overly burdensome. As a starting point, we used data tables from the Jordan Household Expenditure and Income Survey 2017–18 to identify an initial list of consumption items that were likely to comprise the largest share of consumption. Specifically, we included in our initial list the smallest number of items that collectively contributed 90 percent of total consumption. However, recognizing that program participants might differ from the typical Jordanian household, we adjusted this initial list by (1) cross-referencing the consumption survey conducted with refugees in Jordan in 2021 for UNHCR's vulnerability assessment framework, and (2) using

[24] We used the initial list of potential asset index variables identified in the evaluation framework report as a starting point, and excluded durable goods that were extremely common or extremely rare, as well as housing characteristics that showed a weak relationship with self-reported household expenses or had limited variation. We applied the principal components analysis to Cohort 1 and used it to predict the index for Cohort 3. We then divided all participants into asset index quartiles based on the Cohort 1 distribution.

[25] These numbers reflect the sample after we corrected for duplicates across Cohorts 1 and 3 and dropouts from Cohort 3.

[26] These are different from the CEM strata and were intended to broadly ensure the representativeness of the sample and provide Mindset with a primary and replacement sample using an approach they were familiar with from the IGA surveys conducted as part of the DIB.

what we learned during survey pretesting and piloting. These changes to the list included adding, removing, combining, or disaggregating items to make the final list more relevant to the consumption of program participants and minimize respondent burden while maintaining accuracy.

2. Training and implementation

The data collection team, most of whom had previous experience conducting household vulnerability and expenditure surveys, participated in a 3-day in-person training on the survey, including training on protection of human subjects and vulnerable populations, a detailed review of the survey questions, and practice conducting interviews using a series of respondent scenarios. Following training, we then conducted a day of piloting with both Cohort 1 and Cohort 3 participants.

Data collection occurred between late May and late July 2024, a mean of 22.6 months following grant distribution for Cohort 1, and between 1 and 3 months following the beginning of training—and prior to grant disbursement—for Cohort 3. Twenty-three different enumerators conducted the survey in teams of two or three, each with the support of a logistical coordinator who coordinated with potential respondents while in the field. The study used a verbal informed consent statement and procedure approved by the King Hussein Cancer Center Institutional Review Board (IRB) in Jordan and Health Media Lab IRB in the United States. Surveys lasted a mean of almost 50 minutes.

3. Sample sizes and response rates

The response rate for the primary sample was 68 percent for Cohort 1 and 88 percent for Cohort 3. Using replacements, a total of 810 Cohort 1 respondents and 1,056 Cohort 3 respondents completed the survey. We had initially planned to have the same number of respondents (about 930) in both cohorts but adjusted to include more in Cohort 3 when it became clear that we would not reach this target in Cohort 1 given challenges with the response rate and a limited number of replacements given the smaller cohort size.

Achieving a higher response rate for Cohort 1 almost two years after the end of the program using the available contact information proved challenging, despite several measures we took to try to increase it. First, we implemented a systematic tracking effort to reach and update contact information via SMS and phone for Cohort 1 participants at two time points between the 10-month income-generating activity survey and the 23-month impact survey. During the survey, we also worked closely with NEF and *Siraj* centers to reach out to grantees, encourage participation in the survey, and obtain updated information about their availability and willingness to participate. Finally, we conducted a small number of surveys by phone (about 2 percent of the total for Cohort 1) for respondents who had moved to other communities within Jordan or who were unable or unwilling to participate in person.

Table C.1. Final statuses of survey non-respondents

Status	Cohort 1	Cohort 3
Unreachable by phone	176	64
Refused to participate	143	40
Unavailable due to migration, incarceration, or death	36	2
Unreachable or incomplete in the field	24	11
Total incomplete	**379**	**117**

Source: Impact survey

Ultimately, 379 sampled program participants in Cohort 1 were unreachable, or unable or unwilling to participate (**Table C.1**). Of these, 217 were unreachable by phone, in large part due to disconnected phone numbers. 143 participants refused to participate, and an additional 24 were unreachable or incomplete in the field. 33 participants had migrated to other countries, and an additional 4 were incarcerated or deceased. 117 Cohort 3 participants who were sampled initially or added from the replacement list also did not respond to the survey. The reasons were mostly similar to Cohort 1, except that they had lower rates of out-migration. In Appendix E, we assess the implications of non-response for the findings.

D. Additional details on outcomes

The consumption metric is defined as the average monetary value of annual household consumption. **Table C.2** summarizes the data collected and the estimation approach for each category of consumption.

Table C.2. Details on calculation of the consumption metric

Category	Data collected in the impact survey	Calculation approach
Food	– Quantity of food consumed by household members in the previous seven days – The quantities and total costs of food purchased in the prior 30 days	**1.** Calculate typical unit prices for each type of food, survey location and cohort, based on survey data **2.** Apply the unit price to the quantity consumed to estimate the total value of weekly consumption per food item **3.** Sum across all goods and multiply weekly consumption by the number of weeks per year to produce a total annual estimate
Non-food	– Expenditure on transportation, personal care products, utilities, tobacco products, clothing, maintenance of home and vehicles, health care, education, and other miscellaneous items – Reporting periods varied between 1, 3, and 12 months depending on the item and typical consumption patterns in Jordan	**1.** Calculate annualized values by category **2.** Sum across annualized estimates to produce a total annual estimate
Durable goods	– Purchase cost of each durable good owned – Year of purchase – Respondents' estimate of the current market value of the good	**1.** Convert purchase costs to 2024 values using the Jordanian Consumer Price Index **2.** Use the difference between the present value of the purchase cost and the current value to estimate the annual depreciation rate for each type of good **3.** Estimated the value the household derives from the good (intuitively, the cost at which they might be willing to rent it out) as the current value plus interest minus annual depreciation.
Housing	– Rent payments (annual or monthly) – Respondent-estimated cost to rent current home, if owned or used for free	Convert actual or estimated rent to annual values to produce a total annual housing cost

Table C3 summarizes the additional outcomes that we estimated.

Table C3. Definition of additional outcome metrics

Domain	Measures and source
Household savings	Total amount of money household holds in savings
Coping strategies and food security	• The revised Consumption-based Coping Strategy Index (rCSI), measuring strategies to adjust food consumption to bridge limited availability of food in the short term (Maxwell et al. 2008) • The Livelihood-based Coping Strategy Index (World Food Program), measuring broader strategies to meet basic food needs (for example, spending savings, buying food on credit, selling belongings or assets, begging) For both scales, we applied the same severity ratings and weights as a previous study of vulnerable populations in Jordan (REACH 2020).
Household debt	Current amount and types of debt held
Household income and social assistance	Sources and amounts of earned income from wages, businesses, and pensions, and amounts and types of social assistance, including remittances received over the past month
Self-confidence	A modified version of the Rosenberg Self-esteem Scale (Rosenberg 1965), a 10-item scale measuring positive and negative feelings towards the self
Women's social and economic empowerment (women participants only)	• Extent of women's influence in household spending decisions (DHS, Donald et al. 2017) • Women's freedom of movement (Yount et al. 2016), a series of questions about ability to independently visit places (market, doctor, neighbor, etc.)
Sense of safety and well-being	A modified version of Diener's Satisfaction with Life Scale (Diener et al. 1985), a short 5-item instrument designed to measure global cognitive judgments of satisfaction with one's life
School enrollment and attendance	Whether child is enrolled in school and if so, how many days they attended in the previous week
Long-term business metric and financial metrics	A subset of questions from the IGA survey to assess whether grantees were still engaging in IGAs and to measure financial outcomes for grant-supported businesses; administered only to Cohort 1 grantees

E. Analysis approach

We rematched the respondents who completed the survey using CEM based on the characteristics shown in **Table C4** to ensure that the analysis sample remained balanced in socio-demographic characteristics between Cohorts 1 and 3 given small differences in response rates by subgroup. These characteristics were similar to those used in the initial matching before sampling and were selected to optimize baseline balance using CEM while retaining as much of the sample as possible. Out of all survey respondents, 757

out of the 810 respondents in Cohort 1 and 890 out of the 1,056 respondents in Cohort 3 were matched; these comprised the final analysis sample.[27]

We then used regression analysis to estimate impacts by comparing outcomes in the matched treatment and comparison groups by estimating the following regression equation using ordinary least squares:

$$Y_i = \alpha + \beta_1 T_i + \beta_2 X_i + \varepsilon_i$$

where Y_i is the outcome for participant i; T_i is an indicator for treatment, equal to 1 for the treatment group (Cohort 1) and 0 for the comparison group (Cohort 3); and ε_i is a random error term. Although our matching approach minimized pre-existing differences between treatment and comparison groups, we included a set of control variables, X_i, in the model to account for any remaining imbalance between the two groups (Ho et al. 2007; Stuart 2010). X_i consists of indicators for categorical matching variables (respondent gender, refugee status and location at baseline), continuous versions of the remaining matching variables (age, age squared, household size, household size squared, and the continuous asset index), as well as education level (less than secondary, secondary, or more than secondary), which we did not include in the final matching because it led to additional sample size loss without improving overall baseline balance.[28] The coefficient, β_1, provides the regression-adjusted estimate of the impact of the program, which is the difference in outcome means between the treatment and comparison groups after controlling for potential confounders that may contribute to the difference.

Table C4. Matching characteristics for CEM

	Asset index (four quartiles)	
Demographic characteristics	**Housing characteristics**	**Durable goods owned**
– Gender – Refugee status (refugee versus Jordanian) – Youth (25 years or younger versus older than 25 at the time that Cohort 1 was selected) – Program site (five sites) – Household size (3 or less; 4-7; 8 or more)	– Owns home – Number of persons per room	– Car – Computer – Air conditioner – Microwave – Gas/electric heater – Water heater/cooler – Electric fan – Vacuum cleaner – Clothes iron – Television

Notes: Information on matching characteristics was obtained from NEF's vulnerability assessment, which participants complete before entering the program. All matching characteristics were measured as of late 2021, when Cohort 1 was entering the program; NEF obtained some of this information retrospectively for Cohort 3 when this cohort completed the vulnerability assessment in early 2024.

[27] The final analysis sample includes 33 Cohort 3 participants who dropped out of the program after applying and agreeing to participate, but who did not ultimately attend any trainings.

[28] The estimates were not sensitive to including these control variables.

To assess impacts separately for grantees, we rematched using grantee status as an additional CEM matching variable once this became known for Cohort 3 (several months after data collection was completed). For this analysis, the analysis samples comprised 488 Cohort 1 grantees matched to 451 Cohort 3 grantees, and 205 Cohort 1 non-grantees matched to 283 Cohort 3 non-grantees. We then conducted the analysis separately for grantees and non-grantees using the same specification as above

F. Generalizability and internal validity

This Cohort 1 impact analysis sample was broadly similar in characteristics to the full group of Cohort 1 program participants, although the former included slightly greater proportions of women, refugees, adults, and households in the lowest baseline asset quartile.[29] These modest differences were driven by a combination of differential match rates (because Cohort 1 respondents with less common characteristics were less likely to find an exact match in Cohort 3), as well as small differences in survey response rates. The overall similarity in characteristics between the impact analysis sample and the full group of participants for Cohort 1 supports the generalizability of the findings to the latter.

The matching approach successfully resulted in well-balanced Cohort 1 and Cohort 3 analysis samples with similar demographic characteristics and baseline socio-economic characteristics (Appendix Table A3a). The largest differences were that Cohort 3 participants were 5 percentage points less likely to have a disability (because disabilities were rare in both cohorts, but more so in Cohort 3 than Cohort 1) and 6 percentage points more likely to have post-secondary education (the analysis controls for educational attainment, but we did not include it as a matching variable). The estimates were not sensitive to including disability status and post-secondary education in the matching process; this would have ensured balance along these characteristics but led to greater sample size loss and reduced statistical power to detect impacts. Overall, the similarity between the matched Cohort 1 and Cohort 3 samples supports the internal validity of the comparison group design.

Like for the overall analysis sample, the analysis sample for Cohort 1 grantees was broadly similar in characteristics to the full group of Cohort 1 grantees, although it was somewhat more likely to comprise women (7 percentage points) and less likely to comprise youth (8 percentage points) (Appendix Table A2b). The grantee-only analysis sample was also similarly well-balanced in characteristics between Cohort 1 and Cohort 3 (Appendix Table A3b), supporting the internal validity of the grantee-only findings.

[29] To calculate asset quartiles for the IGA survey, we conducted a principal components analysis across all three cohorts based on pre-program housing characteristics and durable goods owned at the time of the vulnerability assessment, which NEF administered to all participants in each cohort shortly before they were selected for the program. We used the coefficients from this analysis to estimate an asset index for each participant. We then divided participants into quartiles by comparing their asset index to the overall distribution across cohorts, with participants in the first quartile being the relatively best off and participants in the fourth quartile being the relatively worst off at the time of selection into the program.

Appendix D: Technical details for the process evaluation

Table D1. Characteristics of the FGD sample for first and second round qualitative data collection

Group	Number of FGD participants	Percent of FGD participants
Location		
Amman	28	17.8%
Irbid	36	22.9%
Kufrsoum	30	19.1%
Russeifa	32	20.4%
Zarqa	31	19.7%
Grant status		
Grantee	120	77.4%
Non-grantee	35	22.6%
Gender		
Women	98	62.4%
Men	59	37.6%
Age		
Adult (26+)	101	65.2%
Youth (18–25)	54	34.8%
Business type		
Home food processing	33	21.3%
Grocery, food trade, etc.	35	22.6%
Trade in clothes and shoes, fabrics, make-up	15	9.7%
Home sewing and tailoring	14	9.0%
Salon, beauty center, gym	10	6.5%
Other	48	31.0%
Total	**157**	**100%**

Notes: Totals are slightly smaller for some breakdowns due to missing data.

Table D2. KII participants for first and second round qualitative data collection

Implementer position	Number of KIIs	Description of role
NEF headquarters staff (repeated in both rounds)		
Program Manager	1	Has overall responsibility for managing project implementation.
Monitoring, Evaluation, and Learning Manager	1	Monitors participants' business projects and outcomes continuously during the training and post-grant period.
Business Development Officer	2	Manages the recruitment and selection of training participants and reviews and scores business plans. Leads post-grant support and business accelerator activities.
Capacity Building Officer	1	Oversees implementation of the training curriculum. Manages the 5-day in-person business skills training and a 2-day online training on gender-based violence. Creates training schedules and oversees logistical arrangements for trainings. Assists with selecting, training, and overseeing trainers.
Field staff at Siraj centers (different centers in each round)		
CBO Head	5	Provides recommendations during selection of trainers, training volunteers, and other staff. Coordinates candidate interviews for these positions. Participates in collective decision-making process for final staff selections.
Trainer	8 (6 individual interviews during Round 1; 2 group interviews during Round 2)	Leads business skills training. Offers direct guidance and tailored support to participants. Conducts initial reviews of grant applications.
Siraj Officer / Senior Siraj Officer	5	Manages logistical arrangements for recruitment, training, and post-grant support. Participates in selection of grant recipients. Conduct data entry as part of monitoring and evaluation efforts.

Appendix E: Sensitivity of impact estimates to survey non-response

Low survey response rates are a challenge in many impact evaluations. (Ghanem et al. 2023; McKenzie 2017), and our study is no exception. Our final impact survey response rates were 68 percent for Cohort 1 (the intervention group) and 88 percent for Cohort 3 (the comparison group), despite our best efforts to increase the response rate for Cohort 1. This differential response rates could potentially introduce bias into our impact estimates by undermining the comparability of the two study groups. Specifically, we might underestimate the DIB's impacts if non-respondents in Cohort 1 are disproportionately "high achievers" (for example, those who were too busy with their successful business to respond to the impact survey). Since their outcomes are excluded from the analysis, the impact estimates could be biased downwards. Conversely, we might overestimate impacts if non-respondents in Cohort 1 are disproportionately "low achievers" (for example, those who were embarrassed that their businesses failed). Our matching approach cannot fully address this potential bias, as non-response may be driven in part by unobserved factors.

To explore the extent to which non-response bias could affect our results, we applied Lee (2009) bounds, which estimates upper and lower bounds impacts under assumptions about the outcomes of non-respondents. This method involves trimming the sample of the study group with the higher response rate (Cohort 3) to match the size of the study group with the lower response rate (Cohort 1). To estimate an upper bound, we exclude observations with the most positive outcomes in Cohort 3, effectively assuming that the additional non-respondents in Cohort 1 are disproportionately "high achievers." Conversely, to estimate a lower bound, we exclude observations with the most negative outcomes in Cohort 3, effectively assuming that the additional non-respondents in Cohort 1 are disproportionately "low achievers.".

Overall, we find that our impact estimates are somewhat sensitive to the Lee bounds assumptions on non-respondents. Table E.1 provides the upper and lower bound impact estimates for three key outcomes in the evaluation: total household consumption, total annual household income, and business ownership, alongside the primary impact estimates from Section VI for comparison. In all cases, the upper bound estimates are more than twice as large as the primary impact estimates, both in the original units of measurement and in standard deviations. For the consumption metric and total annual household income, the lower bound estimates suggest that the DIB program's impacts are effectively zero. The exception is business ownership, where the lower bound still indicates a measurable positive effect.

Table E1. Lee bounds impact estimates

Outcome	Primary impact estimate (for comparison)		Upper bound estimate		Lower bound estimate	
	Units	Standard deviations	Units	Standard deviations	Units	Standard deviations
Consumption metric: Total annual household consumption (JOD)	636**	0.22**	1,524**	1.04**	24	0.01
Total annual household income (JOD)	674**	0.23**	1,739**	1.24**	-141	-0.05
Owns a business (%)	28**	0.60**	49**	1.43**	18**	0.36**

Source: Impact survey.

Notes: Samples include grantees and non-grantees. Comparison means and treatment-comparison differences are estimated using coarsened exact matching weights and regression adjusted with controls for household size and its square, age at the time of program entry for Cohort 1 and its square, the continuous asset index, and education level (less than secondary, secondary, or more than secondary). Lee upper bounds were estimated by excluding the observations with the most positive outcomes from Cohort 3 to match the sample size of Cohort 1 when estimating impacts. Conversely, Lee lower bounds were estimated by excluding the observations with the most negative outcomes in Cohort 3. Effect sizes were calculated using the standard deviation of the comparison group.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

However, these wide bounds represent extreme scenarios and are agnostic as to the likely direction of the bias. To further explore the likely direction of bias, we assessed non-response for Cohort 1 in terms of observed characteristics. Although our matching approach accounts for these characteristics, they might be indicative of unobserved characteristics that affect outcomes. The evidence from this analysis suggests that non-response bias may make it more likely to underestimate impacts, if anything:

- Program data at baseline indicate that survey non-respondents from Cohort 1 were significantly more educated and less likely to belong to the lowest quartile of asset ownership than survey respondents. Because these characteristics are correlated with outcomes like income and consumption, this suggests that non-respondents are more likely to be "high achievers" with more positive outcomes (Table E2). Although this evidence is not definitive, it implies that the exclusion of non-respondents from the analysis would lead us to underestimate program impacts if anything.

- Comparisons in characteristics and outcomes between early and late survey respondents in Cohort 1 lead to a similar conclusion, although the statistical precision is limited due to smaller sample sizes. Specifically, later in the survey data collection period, we made additional efforts to contact non-respondents by coordinating with NEF and Siraj centers to conduct outreach. Those respondents we reached through this additional effort, referred to as "late respondents," might more closely resemble non-respondents (they would themselves have been non-respondents without this additional effort). Late respondents were more likely to own their homes at baseline than early respondents (Table E3). They were also more educated and owned more assets, though the differences between groups are imprecisely estimated. Although late respondents were slightly less likely to own a business, they reported higher total annual consumption and income compared to early respondents, although again these differences are imprecise. These patterns

suggest that the non-respondents—those who ultimately did not respond to the survey—are unlikely to be the "low achievers" based on observable characteristics and outcomes of late respondents. Again, this provides suggestive evidence that the exclusion of non-respondents from the analysis would lead us to underestimate program impacts if anything.

In summary, we conclude that survey non-response poses a challenge to the reliability of our impact estimates; however, the bias is more likely to lead to an underestimate rather than an overestimate of the program's impacts.

Table E2. Comparison of characteristics between Cohort 1 respondents and nonrespondents

Characteristic	Cohort 1 Respondents	Cohort 1 Nonrespondents	Difference
Demographic, household, and grant characteristics:			
Women (%)	86.5	76.8	9.8**
Youth (%)	25.3	40.9	-15.6**
Refugees (%)	34.6	21.9	12.7**
Mean age at the time of program intake (years)	36.4	32.5	3.9**
Head of household (%)	30.9	25.9	5.0
Has a disability (%)	7.9	4.7	3.2*
Literate (%)	96.9	98.7	-1.8*
Mean household size at time of program intake	5.2	5.2	0.0
Education level (%)			
Less than secondary education	31.6	23.5	8.1**
Secondary school	42.7	39.1	3.7
Post-secondary (technical or university)	25.7	37.5	-11.8**
Received a grant (%)	67.9	63.3	4.6
Grant amount (for grantees only, JOD)	565	566	-1
Baseline asset index			
Quartile 1 (%)	28.8	17.9	10.8**
Quartile 2 (%)	23.2	27.2	-4.0
Quartile 3 (%)	24.4	26.4	-1.9
Quartile 4 (%)	23.6	28.5	-4.9
Baseline housing characteristics			
Owns home (%)	29.0	25.6	3.4
Persons per room	1.9	1.8	0.1*
Sample sizes	**757**	**478**	**n.a.**

Source: NEF program activity data and NEF vulnerability assessment.

Notes: n.a. = Not applicable. The asset index is based on housing characteristics and durable goods ownership from NEF's vulnerability assessment, using a principal component analysis for all three cohorts. Youth are defined as being under age 25 at the beginning of the program.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Table E3. Comparison of characteristics and outcomes between Cohort 1 early and late respondents

Characteristic	Cohort 1 early respondents	Cohort 1 late respondents	Difference
Demographic, household, and grant characteristics:			
Women (%)	88.4	76.9	11.5**
Youth (%)	24.1	31.5	-7.4
Refugees (%)	36.3	25.4	10.9*
Mean age at the time of program intake (years)	36.7	35.0	1.7
Head of household (%)	30.6	32.3	-1.7
Has a disability (%)	7.9	7.7	0.2
Literate (%)	96.5	99.2	-2.8**
Mean household size at time of program intake	5.3	5.1	0.2
Education level (%)			
Less than secondary education	31.6	31.5	0.1
Secondary school	43.1	40.8	2.3
Post-secondary (technical or university)	25.3	27.7	-2.4
Received a grant (%)	67.5	70.0	-2.5
Grant amount (for grantees only, JOD)	566.6	556.6	10.0
Baseline asset index			
Quartile 1 (%)	30.1	21.5	8.6*
Quartile 2 (%)	23.1	23.8	-0.8
Quartile 3 (%)	23.1	31.5	-8.5
Quartile 4 (%)	23.7	23.1	0.6
Baseline housing characteristics			
Owns home (%)	27.5	36.9	-9.4*
Persons per room	1.9	1.9	0.0
Outcomes			
Total annual household consumption (JOD)	6,785	7,409	-624
Total annual household income (JOD)	4,367	4,894	-527*
Owns a business (%)	47.4	42.3	5.1*
Sample sizes	**680**	**130**	**n.a.**

Source: NEF program activity data, NEF vulnerability assessment, and impact survey (outcome only).

Notes: n.a. = Not applicable. The asset index is based on housing characteristics and durable goods ownership from NEF's vulnerability assessment, using a principal component analysis for all three cohorts. Youth are defined as being under age 25 at the beginning of the program.

*/** Difference significantly different from zero at the .05/.01 levels, two-tailed test using heteroskedasticity-robust standard errors.

Appendix F. Evidence from reference studies used to benchmark the consumption metric

Table F.1. Overview of relevant studies used to benchmark the consumption metric

Country	Program name	Program description	Impacts on enterprise profits (SDs)	Impacts on income (SDs)	Impacts on household consumption (SDs)
Uganda	Youth Opportunities Program (YOP)	Group training and business start-up support, whereby groups of youth in conflict-affected areas applied for grants averaging $382 per person to start nonagricultural businesses.	**0.16** (2 years) **0.17** (4 years)	*Not available*	**0.18** (4 years)
Liberia	Action on Armed Violence (AoAV)	Residential coursework for male former fighters in agriculture and animal husbandry, along with counseling, life skills classes, and a package of tools and supplies to start business, valued at $125.	**No effect after 14 months**	**0.12** (14 months)	*Not available*
Sri Lanka	Start and Improve Your Business (SIYB)	A widely used entrepreneurship training course from the International Labor Organization that helps participants select and operationalize feasible business ideas. Training was for urban women and included cash grants of $129.	**0.16** (7-8 months) **0.18** (15-16 months) **No effect after 2 years** **No effect for potential new business owners at any time point**	**No effect after 7, 15, and 24 months**	*Not available*
Uganda	Start and Improve Your Business (SIYB)	Similar to the program in Sri Lanka (above), with cash grants of $200.	**No effect after 9 months**	*Not available*	**No effect after 9 months**
Uganda	AVSI Women's Income Generating Support (WINGS)	Business skills training for rural women, cash grants of $150 to support businesses, and ongoing support through supervision and self-help group formation.	*Not available*	**0.30** (16 months)	**0.38** (16 months)
Bangladesh	BRAC's Targeting the Ultra Poor (TUP) Program	Intensive skills training for rural women related to livestock businesses. Participants chose among six livestock packages, to the value of $140.	*Not available*	**0.15** (2 years) **0.17** (4 years)	**0.07** (2 years) **0.10** (4 years)

Source: Keaveney et al. 2018

Appendix G: Cost-effectiveness

In this appendix we describe the main program costs, provide details on the methodology of the cost effectiveness analysis, and discuss the findings in terms of overall and per-beneficiary costs and benefits (**Box G.1**).

A. Estimating costs

We collected detailed data from NEF on realized program costs using the ingredients method, which entails categorizing expenses by cost type and specifying the years in which these costs were incurred. Our focus was limited to cost data related to inputs essential for replicating the program in other contexts, and excludes DIB-specific costs that are not strictly required for replication.

A summary of key costs, as well as those excluded, summed across years is presented in **Table G.1.**

Box G.1. Costs and benefits for whom?

Because we found no statistically significant impacts on non-grantees, our cost effectiveness analysis focuses on **benefits for the 3,416 program grantees** and assumes zero benefits for non-grantees. Similarly, we use the number of grantees when reporting per-beneficiary costs.

However, our estimates of **program costs reflect the costs of the program overall.** These include the cost of providing business development trainings for an additional 2,200 participants who were not selected for grants as well as collecting vulnerability assessment data from more than 4,000 additional potential participants who were not selected for the program. As discussed in Section VII, program implementers perceived oversubscription combined with a rigorous approach to grant selection as one of the keys to the overall success of the program. Thus, these costs should be considered a core part of program implementation and critical for replication.

Table G.1. Summary of total costs for the DIB program (simple sum across years)

Standard cost category	Total nominal costs	Description
Implementation costs	$6,265,722	Costs for providing financial and non-financial services to program participants through the identification and recruitment, capacity-building, and post-grant support phases. They include cost associated with the following specific activities: / Financial services to grantees, including both the grants and the advanced grants provided to select grantees / Local transportation for local team members who were engaged in activities such as coaching, training follow-ups, outreach, and data collection / Internet access for delivering coaching and training remotely / Monitoring, evaluation, and learning (MEL) related activities and data systems for tracking and assessing the effectiveness and delivery of non-financial services / Support for partner CBOs, including needs assessments, capacity building, and recruitment of implementation staff / Social inclusion & protection programming, such as direct social assistance to beneficiaries and childcare support

Standard cost category	Total nominal costs	Description
		/ A small number of events, including community outreach and market access events
Facility costs	$267,805	Costs associated with rent and utilities for program-related spaces and furniture, office supplies, and equipment purchased for the program, such as laptops, printers, and mobile devices.
Staffing costs	$1,369,879	The value of time charged directly by program staff, as well as essential staff-intensive professional services such as legal, payroll, consulting, and recruiting services. Additionally, non-staffing costs such as postage and delivery expenses, compliance services, and taxes and licenses are also reported here because they could not be separated from these services based on the cost data NEF provided.
Administrative costs	$1,018,595	Overhead costs are shared costs not specific to the program but are necessary for NEF to operate effectively, and which are allocated proportionately across programs operated by NEF. They include rent for NEF headquarters , utilities, administrative staffing, and general office expenses.
Other costs	$196,654	Other costs include international travel for international and Jordan-based staff. Travel was undertaken for a combination of administrative and implementation-related purposes including program support and oversight, MEL activities, stakeholder engagement, event facilitation, and capacity building, and could not be allocated to other, more specific ingredients based on the cost data provided by NEF.
Total nominal costs	**$9,118,656**	**n/a**
Excluded costs	$244,892	Excluded costs are those reported by NEF as specific to managing the DIB. Since any future implementation might not be DIB-funded, these costs are not considered core program costs critical to replication.

Source: NEF

We adjusted these costs by converting the costs reported in each year to 2024 dollars and applying a discount rate of 12 percent per annum to account for the time value of money and enable direct comparisons with the benefit streams estimated later. The adjusted total estimated program cost was $10.2 million in 2024 dollars, or about $3,000 per grantee. More information on the approach to discounting is provided in the following section.

Table G.2 provides additional details on the activities associated with the implementation phase and how responsibilities were allocated to NEF and CBO staff. This allocation has important implications for the costs of any future replication or scale-up.

Table G.2. Detailed overview of roles and responsibilities associated with implementation costs

Phase	Activities included in each phase	NEF responsibilities	CBO responsibilities
Relevant staff	n/a	– Business development specialists – Capacity Building & Training Specialist – Senior Siraj Officers & Siraj Officers – MEL Manager & Assistant	– Head of CBO – Volunteers* – Master Trainers – Financial Officer, – Secretary – Board Members* * denotes participation only in identification and recruitment
Identification and recruitment	– Conducted door-to-door canvassing – Conducted vulnerability assessments and preliminary interviews	– Provided technical expertise and co-developed curricula – Managed and deployed of digital training platforms; – Provided oversight of MEL systems – Conducted strategic planning and coordination of activities	– Led community outreach and engagement – Provided logistical arrangements for activities – Facilitated connections with local stakeholders and institutions – Promoted awareness and participation within the community
Training and capacity building	– Provided digital and in-person trainings to trainees – Conducted technical and vocational skills building for selected grantees – Provided virtual, and in-person coaching sessions to small groups	– Provided technical oversight of master trainers – Co-designed blended (digital and in-person) delivery of training content – Established MEL frameworks – Conducted capacity building of CBOs	– Delivered business development training and coaching sessions (master trainers) – Mobilized participants and coordinated venue setup and materials
Post-grant support	– Provided 1:1 in-person and digital business support and mentorship to grantees	– Provided ongoing technical assistance; – Ensured quality assurance across M&E tools – Guided strategic adaptation and improvement of activities	– Provided support and mentorship (master trainers) – Logistical support for master trainers – Contributed to adaptive program management through regular feedback and reporting – Co-led reflection sessions to refine approaches based on field realities

Source: NEF

Note: Appendix Table D2 summarizes the specific roles and responsibilities of Jordan-based NEF staff and CBO leadership.

B. Estimating benefits

To assess the benefits of the DIB program, we focused on two primary benefit streams, as estimated from the evaluation:[30]

1. **Monthly business profits**: Using business profits to value benefits is consistent with NEF's ex-ante calculation of the return on investment (ROI) for the DIB program, which was used to justify the financial investment in the program. Estimating the ex-post ROI using this measure is therefore of primary interest to the DIB parties because it enables them to assess the realized financial returns of the program. We calculated these benefits for all cohorts, using cohort-specific estimates of profits where available and information about the number of grantees per cohort.

2. **Annual household consumption**, While using business profits to measure benefits provides insight into the financial rate of return from the investor's perspective, the impact on household consumption at the 23-month mark offers a more comprehensive measure of its net economic returns to society. Although we recognize that impacts on annual household consumption, like other measures, may not fully capture the program's benefits, this measure offers several advantages. First, increasing household consumption was a key outcome in the program logic that underpinned the evaluation design. The impact survey was designed specifically to measure annual household consumption, adhering closely to established best practices in the literature for precise measurement (Deaton and Zaidi 2002). Second, the evaluation was designed specifically to attribute impacts on consumption to the program by comparing outcomes between the intervention group (Cohort 1 participants) and a carefully selected comparison group (matched Cohort 3 participants). Third, although the impact survey only measured consumption at one point in time, consumption is less subject to fluctuations over time than other measures because households tend to smooth income fluctuations in their consumption habits (Deaton and Zaidi 2002). In contrast, both the IGA and impacts surveys only captured business profits in the past month, which may not accurately reflect annual profits due to seasonal fluctuations in monthly profits. Fourth, consumption is one of the most common indicators of household well-being in the literature and in other cost effectiveness analyses of livelihoods programs (Sulaiman et al. 2016). Using impacts on household consumption to value benefits in the analysis facilitates more meaningful comparisons of benefit-cost ratios across programs. which offers a more comprehensive measure of its net economic returns to society.[31] The initial impact on consumption was based on our estimate of the program's impact on annual household consumption in the second year after program participation ($1,332 for grantees). We then applied the same impact estimates for Cohort 1 to all cohorts, multiplying it by the number of grantees per cohort.

Both costs and benefits accrued or will accrue over time, but over different time periods (costs in the past and benefits in the past, present and future), and it is important to account as accurately as possible for

[30] The two streams are overlapping and alternative approaches to thinking about program benefits, not mutually exclusive measures that can be added together without significant double-counting.

[31] Although increases in household consumption by program beneficiaries might in theory underestimate net societal benefits because increases in business profits are offset by decreases in humanitarian assistance received, our evaluation found only small impacts on the receipt of other assistance.

how the value of money changes over time due to inflation and the opportunity cost of alternative investments. It is also necessary to make assumptions about how business survival and growth will evolve in the future, beyond the period covered by the evaluation. The parameters we include in the estimated benefit streams are as follows (we discuss the specific parameter assumptions below):

1. We select a **time horizon,** or the maximum number of years over which we are confident that effects will persist and be attributable to the program. Factors that are likely to erode impacts over time include loss of business knowledge and skills after the end of training, retirement by older beneficiaries, and workforce transitions by younger ones.

2. For **business profits**, we use the estimated mean business profits for active businesses at 10 months post-grant as our base estimate. For **consumption,** we use the estimated impact on the annual household consumption of grantees at 23 months. We adjust each benefit type for each cohort and year based on the assumed business failure rate and business growth rates, as described below.

    - We prorated both benefit streams by an estimated **business failure rate** in each year, reflecting the proportion of businesses no longer operating.[32] The business failure rate directly affects estimated benefits because we assume that businesses that stop operating yield zero future profits and proportionally reduced consumption benefits.

    - The calculations also incorporate a **business growth rate** to reflect the projected change in business profits over time, which we assume is reflected in a proportional change in impacts on household consumption (due to increased business income). This could either be positive or negative; surviving businesses could either expand and grow, or decline if grantees struggle or become busier with other economic and non-economic activities.

3. We apply **a Jordan-specific inflation adjustment** to bring all dollar values into 2024 dollars, the year when the impact survey was conducted. This ensures consistency with the cost estimates and simplifies calculations and interpretation.

4. We then apply a **discount rate** to reflect the time value of money and the opportunity costs of making alternative investments, so that benefits (and costs) incurred in earlier years are valued more than those incurred in later years, We used the same discount rate for benefits and costs.

To select the values to use in these adjustments, where not already available in the data, we reviewed guidance on conducting benefit-cost analyses, including the standard approaches used by aid organizations and multilateral development bank, as well as literature on small and micro-enterprises in low-and middle-income countries to determine a plausible "base scenario" for the analysis. Estimates and parameters used and the rationale for each, are summarized in **Table G.3**. In Section E below, we describe the approach to testing the sensitivity of our findings to these parameters and the results of those tests.

[32] When grants were awarded in the middle of a calendar year, we also prorated the benefits for that year to begin accruing on the median grant award date.

Table G.3. Parameters used to estimate benefits in the base scenario

Parameter and value	Rationale
Time horizon: 2034, or 10 years from the final year of costs (2025)	Although we originally proposed computing benefits over 20 years, a further review indicates that 20 years may be too long given uncertainties about the duration of effects for livelihoods programming. Moreover, Campos et al. (2024) finds that effects of an entrepreneurship program on trainees in Togo persisted after at least 8 years, but other, similar programs show shorter-term effects (Frese 2024). We selected 10 years as a "middle ground" between alternative scenarios of 5, 10, and 20 years. In section E below, we discuss the sensitivity of results to this parameter.
Business profits: • **Year 1:** Cohort-specific mean values from IGA survey • **Thereafter:** Adjusted by multiplying by business failure and growth rates	The IGA survey provides cohort-specific estimates of mean business profits in year 1 (primarily based on written records). Our adjustments for other years reflect that profits are zero for businesses that fail, while our business growth parameter is intended to be interpreted as growth in profits.
Consumption impact: • **Year 1:** $1,332 from the impact estimate, adjusted for inflation and to reflect a (lower) measured business failure rate • **Year 2:** $1,332 • **Thereafter:** Adjusted by multiplying by business failure and growth rates	The impact estimate applies to Cohort 1 grantees in year 2, reflecting a 24% business failure rate at that point. We applied this year 2 estimate to grantees in Cohorts 2 and 3 too, because cohort-specific estimates were not available. Our adjustments for other years based on business failure and growth rates reflect the evaluation finding that consumption impacts are driven by increased business income.
Business failure rate: • **Year 1:** Cohort-specific values from IGA survey • **Year 2:** 24% • **Thereafter:** 20% annually	The impact survey conducted at 23 months post-grant showed that 76 percent of Cohort 1 grantees satisfied the requirements of the business metric. Moreover, the results of a panel data analysis from Vietnam showed approximately 39 percent of household businesses survived after 5 years (Vijverberg and Haughton 2002). An average annual business failure rate of 20 percent applied after year 2 replicates this finding.
Business growth rate (for surviving businesses): 0%	The Cohort 1 impact survey, which was conducted approximately 13 months after the IGA survey, showed a decline in average profits. However, this might reflect differences in measurement; reported profits in the impact survey were based mostly on self-reports, while in the IGA survey they were mostly based on a review of business documents. It is therefore difficult to estimate a precise business growth rate from available data on profits. However, we did find that take-home business income for Cohort 1 grantees (self-reported in both rounds) was almost identical at 10 months and 23 months. Moreover, in a study of household businesses in Vietnam, McCaig and Pavcnik (2017) found a largely flat relationship between revenue for household businesses and business age over 20 years. Together, this suggests that it is reasonable to assume a flat growth rate in our base scenario.
Inflation adjustments: Varies by year between 1.6-4.3%	We applied inflation rates provided by the Jordanian Bureau of Statistics to convert all measured values to 2024 dollars (the year when the impact survey was conducted), prior to discounting.

Parameter and value	Rationale
Discount rate: 12%	In developing countries, discount rates typically range from 8 to 15 percent (Zhuang, Juzhong et al. 2007). Several aid organizations and multilateral development banks apply rates of 10 percent or higher (USAID 2022).[33] Based on these ranges, a 12 percent discount is a reasonable choice for evaluating the costs and benefits of the DIB program.

C. Results: Base Scenario

We applied the parameters summarized above to determine the net present value of benefits. Results are broadly similar across both benefit streams, and point to strongly positive net benefits.

- **Business profits.** We estimate that over 10 years, surviving businesses will generate approximately $20.1 million in business profits, or $9.9 million in net profits after subtracting program costs (**Figure G.1**). This translates into net benefits of $2,900 per grantee. The benefit-cost ratio is 1.98 meaning that the program generated $1.98 in net profits for every dollar invested.[34]
- **Household consumption:** Over 10 years, we estimate that the project would generate $22.0 million in additional household consumption, or $11.8 million in net benefits after subtracting program costs. This translates into net benefits of $3,500 per grantee, and a benefit-cost ratio of 2.16, meaning that the program generated $2.16 in net consumption for every dollar invested.

[33] The following are the discount rates used by several aid organizations and multilateral development banks: Millennium Challenge Corporation (10 percent); United States Agency for International Development (12 percent); African Development Bank (12 percent); Asian Development Bank (9 percent); and Inter-American Development Bank (12 percent).

[34] NEF also used business revenues as an alternative measure of benefits in its ex-ante calculations, given that business profits might be negatively affected by the cost of investments in new staff, equipment, and other fixed assets, which are desirable for early-stage businesses. For completeness, we also conducted calculations using measure of benefits, and found a benefit-cost ratio of 5.11. However, we consider this estimate less appropriate given that the evaluation measured monthly revenues 10 months after grants were disbursed—that is, after many up-front investments would have been completed—and because it risks classifying loss-making businesses as a good investment.

Figure G.1a. Total costs and estimated benefits over 10 years

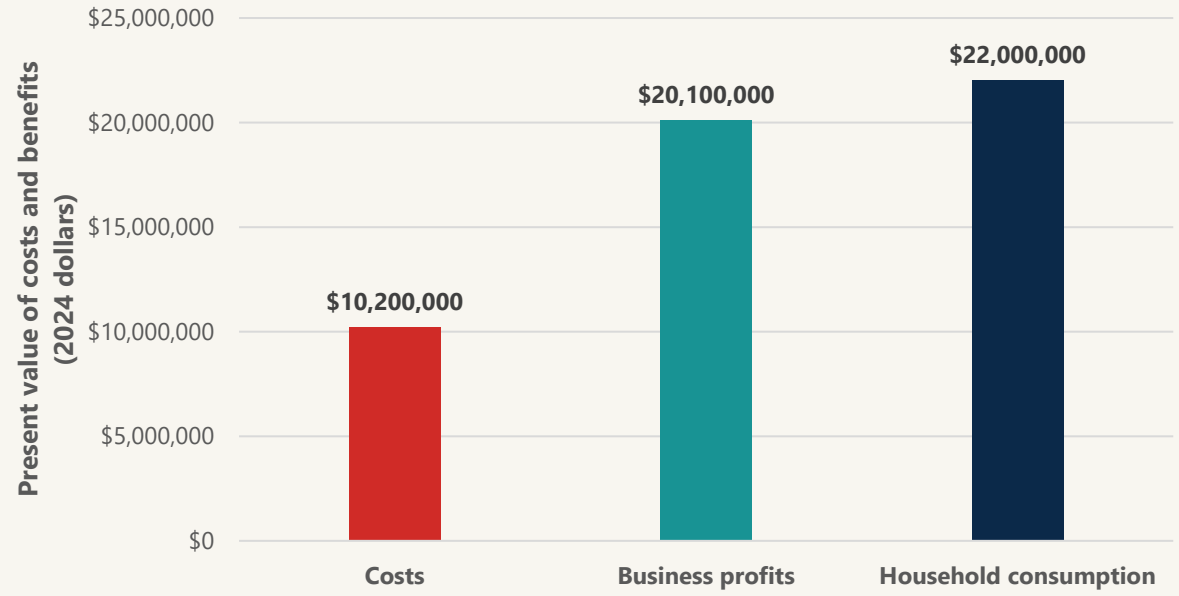
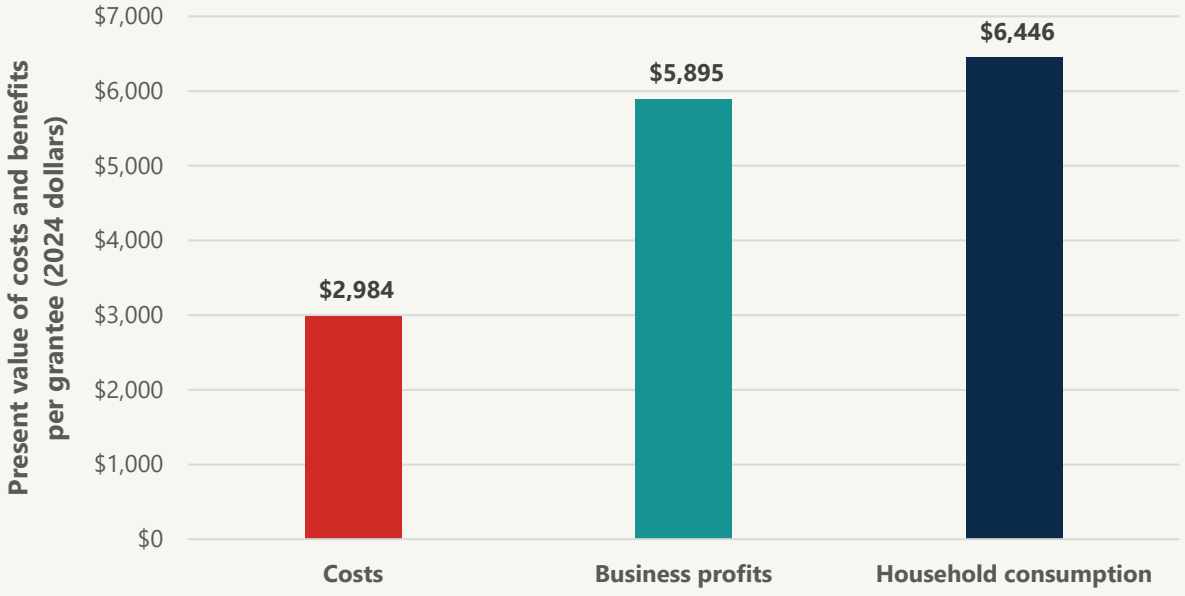


Figure G.1b. Per-grantee costs and estimated benefits over 10 years



D. Sensitivity analysis

We conducted a Monte Carlo simulation that randomly and independently varied the two key unknown parameters: business survival and business growth for surviving businesses. In other words, we used randomly generated parameters for business survival and growth within a specified range to estimate 10,000 different scenarios and calculated the resulting benefit streams to generate a benefit-cost ratio for each scenario. The range of values used for each parameter and the rationale for each are as follows:

- **The random values for business failure rates ranged between 10 percent and 50 percent.** A 10 percent failure rate reflects the most generous interpretation of the impact survey results where 86 percent of grantees self-reported having an active business after 23 months (before applying the DIB business metric definition), compared to 98.5 percent 13 months earlier. A 50 percent failure replicates recent reports that 90 percent of new businesses (including in the MENA region) fail, typically within 3 years (Al-Yahya and Airey 2013).
- **We allowed annual business growth for surviving businesses to vary between -10 percent and 10** percent to, without a strong long-term evidence base, reflect a wide range of plausible outcomes.

We find that the program's overall cost-effectiveness under a wide range of scenarios is positive and robust even to more extreme negative assumptions. The benefit-cost ratio for business profits falls between 1.33 and 2.49, with a median of 1.72 (**Figure G.2a**). The benefit-cost ratio for household consumption ranges between 1.45 and 2.73, with a median of 1.88 (**Figure G.2b**).

Figure G.2a. Distribution of benefit-cost ratios for profits

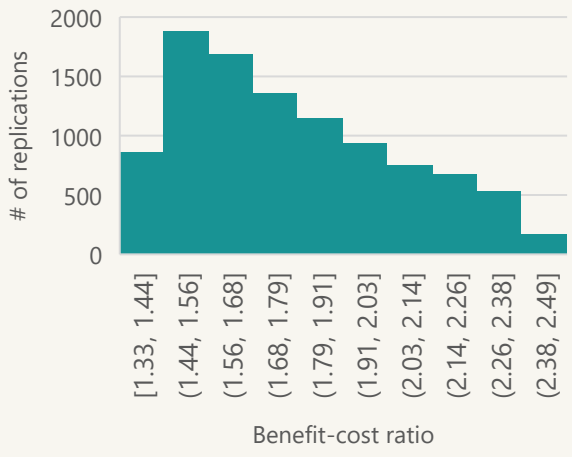
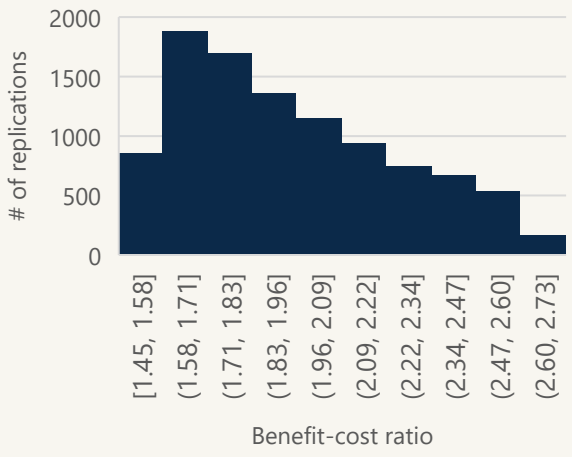


Figure G.2b. Distribution of benefit-cost ratios for household consumption



Additionally, we compared the results of this simulation over a 5, 10, and 15 year time horizon and with alternative discount rates, and find that the results are not sensitive to these assumptions. Reducing the time horizon to a conservative 5 years while leaving the other base assumptions as-is does

not significantly reduce minimum or median values of the benefit-cost ratio for either profits (minimum 1.34 and median 1.62) or consumption (minimum 1.47 and median 1.78). Instead, a time horizon of 5 years reduces the likelihood of more extreme positive results found with a longer horizon, such as benefit-cost ratios above 2.20. There was little difference between 10 and 15 years because limited benefits are accrued after 10 years when most businesses are projected to have ceased operating and remaining benefit flows are severely discounted. Similarly, increasing the discount rate to 15 percent but leaving the other base assumptions unchanged only modestly reduces the benefit-cost ratio for profits and consumption to 1.84 and 2.01, respectively, while reducing it to 10 percent increases the ratios to 2.08 and 2.27, respectively.

E. Comparing cost effectiveness with related programs

Finally, we compare the findings to the cost-effectiveness of comparable programs in low and middle-income countries. Sulaiman et al. (2016) compare the cost-effectiveness of three kinds of social protection interventions in low- and middle-income countries,, none of which were in the MENA region : (1) livelihoods programs (30 studies, primarily in rural areas and with an agriculture focus), (2) unconditional cash transfer programs (11 studies, several with a micro-entrepreneurship focus), and (2) poverty graduation programs that often combine livelihoods programming with asset or cash transfers (7 studies). The review includes only studies that measured impacts of these three groups of interventions on household consumption and/or income, which enables direct comparisons to our findings on cost-effectiveness in terms of consumption. The DIB program overlaps with all three kinds of program in that it provided both livelihoods training and cash grants. Thus, it is worth drawing comparisons to each of them.

To enable comparisons to reported cost effectiveness measures in the studies included in Sulaiman et al. (2016), we calculated an alternative simple benefit-cost ratio that takes a measured point-in-time impact on annual consumption (that is, the impact on annualized household consumption at the time of the follow-up survey, which varies across studies) and divides it by the program costs per beneficiary. For example, a ratio of 0.1 indicates that the impact on annual consumption is 10 cents for every dollar spent per household for the program, and the effects would need to persist for 10 years for costs and benefits to equalize, ignoring discounting. The DIB program's simple benefit-cost ratio calculated using this approach is 0.49, meaning that the effects would need to persist for only two years to exceed the costs per beneficiary.

At a cost of nearly $3,000 per grantee, the DIB program had higher costs per beneficiary than the other livelihoods programs cited by Sulaiman et al. (2016), which cost $1,147 per beneficiary on average (a range of between $236 and $3,700). Cash transfers averaged only about $232 per beneficiary, plus unreported administrative costs, while the average cost of included graduation programs was $1,148 per beneficiary. Despite higher costs, the simple benefit-cost ratio of 0.49 for the DIB program is larger than for the livelihood programs included in the study, which had a mean simple benefit-cost ratio of 0.20. The equivalent mean ratios for cash transfers and graduation programs were 0.29 and 0.11, respectively.

Estimates of program benefits are often based on measurement of outcomes in the short run (less than one year), and there is mixed evidence on the extent to which these benefits persist over time. Estimated benefits and cost-effectiveness may be overstated if program effects fade. Therefore, it is also useful to compare our cost effectiveness estimates for consumption, which are based on benefits measured two

years post-grant, to other programs in Sulaiman et al. (2016) where benefits were measured beyond one year. With this restriction, the DIB program compares even more favorably to other programs. For those other programs with longer-term measures of benefits, the mean simple benefit-cost ratio was only 0.09 for livelihoods programs and 0.11 for cash transfer programs. Our cost-effectiveness findings are most similar to those from a micro entrepreneurship support program in Santiago, Chile, which similar to the DIB program, provided in-kind start-up capital of about $600 and 60 hours of business practices training in an urban environment (Martinez et al. 2016). This project had a simple benefit-cost ratio of 0.50, which is nearly identical to our estimated ratio of 0.49. This suggests that offering micro-entrepreneurship training and sizeable cash grants to urban populations is a promising and cost-effective intervention approach across diverse country contexts.

Mathematica Inc.

Our employee-owners work nationwide and around the world.

Find us at **mathematica.org** and **edi-global.com**.