DECENTRALIZED EVALUATION FOR EVIDENCE-BASED DECISION MAKING

SAVING LIVES
CHANGING LIVES

Decentralized Evaluation

Support for Strengthening Resilience of Vulnerable Groups in Ethiopia: The Fresh Food Voucher Programme Expansion in Amhara Region

January 2018 to December 2020

Evaluation Report (February 2021)

February 2021
WFP Country Office Ethiopia
Evaluation Manager: Dr. Alexandra Priebe

Prepared by
Prof. Dr. Markus Frölich, Team Leader (C4ED)
Dr. Giulia Montresor, Evaluator (C4ED)
Dr. Elazar Tadesse, Evaluator (C4ED)
Ariane Gordan, Research Manager (C4ED)

WFP
World Food Programme

Acknowledgements

We thank WFP for commissioning this study and for the financial and technical support without which this evaluation would not have been possible. We are grateful to WFP Ethiopia (Dr. Alexandra Priebe, Dr. Filippo Dibari, Emma Lefu, Getachew Damtew, Habtamu Amsalu, Andualem Agiz, Mamo Getahun, Zebibia Ayenew, Tadesse Behre, Ambachew Tesfaye) and the WFP Regional Office (Gabrielle Tremblay, Nikki Zimmerman) for their guidance and support. We further thank the ERG members for their time and helpful comments.

We thank the Horn Economic & Social Policy Institute (Abdurohman Hussien, Fentahun Lakew and Awnu Dino), as well as their team of 15 enumerators for their critical role in the fieldwork and the collection of high-quality survey data. We further thank the qualitative field researchers (Nolawit Teshome, Hargewoyin Habtegebriel, Dr. Rukya Hassen), as well as the transcribers, translators and note takers for their role in the qualitative data collection and analysis.

Further, we thank Dr. Robert Poppe, Dr. Thomas Eekhout, Laura Peters, Peter Derheim and Veronica Notaro for their exceptional work provided at different stages of the evaluation.

Finally, we would like to express deep gratitude to the woreda officials for giving us approval to enter their communities; to the health extension workers for their support in setting up the phone interviews with beneficiaries; and to the respondents for their time during the interviews.

Disclaimer

Table of Contents

List of Tables

Table of Figures

Executive Summary

This report presents the findings of the end-line activity evaluation of the Fresh Food Voucher (FFV) Programme Expansion in Amhara Region between 2018 and 2020. The FFV Programme Expansion in Amhara Region resulted from the successful pilot launched in three woredas (Habru, Raya Kobo, and Dessie Zuria) in 2017 and was expanded in January 2018 to an additional four woredas (Dawa Chefa, Kalu, Seqota and Mekdela) and aimed to cover 27,000 households in total for both phases. The expansion was funded by a 7 million Euro grant from BMZ/KfW. This evaluation is commissioned by the World Food Programme (WFP) Ethiopia Country Office (ETHCO) and covers the period from January 2018 to December 2020.

This end-line evaluation follows two objectives: accountability and learning. It shall inform on the performance and results of the FFV Programme Expansion within the Seqota Declaration and/or rural national social protection programme (PSNP) geographical area and wherever the same programme outcome may be relevant in the future. This evaluation will inform why certain results occurred and advise on good practices and lessons for future programming. The evaluation was conducted in the midst of the COVID-19 pandemic and accordingly faced various challenges. The expected users for this Evaluation Report are WFP and BMZ/ KfW as the main donors, as well as the Government of Ethiopia including the Seqota Declaration and its affiliated ministries, and other donors and implementers involved in nutrition programming.

Ethiopia in general, and Amhara Region in particular, still have high rates of malnutrition including stunting. Dietary diversity is low for children between 6 and 23 months as well as their mothers. The "Fill the Nutrient Gap" Analysis in 2020 found that 3 out of 4 households could not afford a diet that meets all nutrient requirements, with breastfeeding women and their children among those with the highest risk of an inadequate diet. The FFV programme's main objectives are to 1) increase the dietary diversity of pregnant and lactating women (PLW) and children under two years, 2) stimulate the local fresh food market and 3) strengthen social and behavioural change to generate more demand for fresh food in seven woredas in the Amhara Region. To reach these objectives, the programme includes a three-pronged approach to availability, access, demand and utilization of fresh foods: the main innovation of the programme are mobile money vouchers to beneficiaries redeemable for vegetables, fruits and animal source foods to enhance access to fresh foods. The voucher is meant to be a 'top-up' of transfers granted by the PSNP to fill the gap between the affordability and the cost of a nutritious diet. However, PSNP-eligible households not enrolled in PSNP were registered in the pilot. A national service provider provided the transfer. The FFV transfer component is coupled with retailer engagement and trainings for fresh food suppliers and retailers in rural markets, as well as Social and Behaviour Change Communication (SBCC) activities.

Methodology

The evaluation was designed to assess the Fresh Food Voucher Programme Expansion in Amhara Region between 2018 and 2020 against the following evaluation criteria: relevance, effectiveness and efficiency, and sustainability. The main evaluation questions (EQ) were:

Relevance:

- *Q1.1: Is the project aligned with the national government's policies and strategies to support the reduction of stunting via increased diet diversity?*
- *Q1.2: How relevant has the approach been to beneficiaries?*
- *Q1.3: How did the programme interventions adapt to the COVID-19 pandemic?*

Effectiveness:

- *Q2.1: Did the project reach the intended beneficiaries with the right mix of assistance?*
- *Q2.2: What are the effects of the project on availability of fresh foods?*
- *Q2.3: What are the effects of the project on access to fresh foods?*
- *Q2.4: What are the effects of the project on the demand for fresh foods?*
- *Q2.5: Did the intervention produce the expected nutritional results?*
- *Q2.6: Is there an increase of financial inclusion and financial autonomy among female beneficiaries?*
- *Q2.7: Was the intervention efficient compared to possible alternatives?*

Sustainability:

- *Q3.1: What is the government readiness to take over the programme?*
- *Q3.2: Is the programme capable of overcoming future challenges and bottlenecks?*
- *Q3.3: Will behavioural changes related to nutrition last after the programme?*

In order to respond to these questions, the evaluation team conducted a mixed-methods evaluation approach using quantitative and qualitative data sources. Evaluation questions on relevance are covered through qualitative analysis. To assess the effectiveness and efficiency of the programme, we use both quantitative and qualitative tools and triangulate the findings. We cross-check certain information collected from household interviews with data from M&E systems. The evaluation was not able to make any assessments of the causal impact of the intervention, as no credible counterfactual could be established. Instead, we use descriptive statistics of key output and outcome indicators to gather insights into the successes and implementation problems of the programme.

The main limitations included the COVID-19 pandemic, as well as the short study period and the Tigray conflict bordering one of the study areas (Seqota woreda). The COVID-19 pandemic implied that all data collection had to rely on remote means, that is remote phone-based data collection had to be used for both quantitative and qualitative interviews. Measures were taken to mitigate against these limitations as far as possible, including dropping Seqota woreda from the sample and engagement of woreda officials and Health Extension Workers (HEW) before the start of the data collection in the verification of phone numbers and reaching out to respondents beforehand to ensure awareness of the upcoming survey. Nevertheless, the reliability and richness of the collected data via short phone interviews is clearly limited.

Key Findings and Conclusions: The key findings and conclusions of the evaluation team are summarised below, structured according to the main evaluation criteria.

Relevance: Relevance of the programme was assessed in terms of alignment into national nutrition policy, usability and appropriateness perceived by beneficiaries, and adaptation to the COVID-19 pandemic. We found that the FFV is closely aligned with key recommendations and strategies of national nutrition policies, such as the National Nutrition Strategy (NNS)/Programme, the Seqota Declaration, and the PSNP. The main aspects highlighted are the importance of access to nutritious fresh foods such as fruit, vegetables and eggs, demand creation and awareness-raising with pregnant

and lactating women and caregivers of children under two years, as well as the strengthening of nutrient-dense foods at local markets.

Beneficiaries largely welcomed the voucher as a means to purchase fresh foods that they otherwise could not afford. A main concern was voiced in the amount of the voucher received, which often seemed to be perceived as too little to purchase all fresh food groups for the entire family. Combined with information emerging from other evaluation questions, a likely reason for this inadequacy may be that the FFV was meant to be a top-up to the PSNP transfer, but many FFV beneficiaries did not receive PSNP support. Thus, while subsidizing fresh food purchases was considered highly relevant to fulfil the needs of a nutrient-dense diet, the full potential could not be realised given the imperfect complementarity between the PSNP and FFV.

With the lockdowns and other measures due to COVID-19 and resulting restrictions in interactions, the programme had to be adapted in order to still be able to reach its beneficiaries. This mainly had consequences in the interactive components such as market interactions, support and interpersonal SBCC activities. While any interactive activities had to be suspended, the programme reacted quickly with the development and roll-out of additional mobile awareness-raising campaigns through SMS and phone calls.

Effectiveness: To determine if the programme worked as intended, it was first analysed if the intended target group was reached with the intended mix of assistance (vouchers and SBCC). Enrolled beneficiaries did largely demonstrate the main eligibility criteria, i.e., households with pregnant or lactating women or children below 2 years of age. However, only 20% of beneficiaries had received a PSNP transfer in the previous year. This implies that for the majority of FFV beneficiaries, the voucher did not serve as a top-up to the PSNP support as intended by the programme, with potential consequences to the adequacy of the voucher amount. Further, we found that a large fraction of those recently graduated from the programme (between October and December 2020) would still have been eligible (according to their reported household structure) in terms of having a pregnant or lactating woman or children under 2 in the household. For about 62% of those graduated households, this may have happened as the child for which registration had taken place became older than 2 and another pregnancy could not be supported by the programme due to capacity constraints. For the remainder, however, this was not the case as no older children have been reported in the household. No clear conclusions could be reached for this large share of eligible but graduated households, partly also because of limited verifiability due to inaccessibility because of COVID-19.

A major bottleneck was the limited reliability of the mobile phone network, which continually affected the implementation of the voucher programme. While the FFV transfers were intended to be disbursed every month, this has only happened for about 60% of beneficiaries. This discrepancy has been explained by network or other technical issues, and missed transfers were disbursed in the following months. However, these delays still make it difficult for beneficiaries and traders to plan ahead for purchases as well as supplies of fresh foods. In terms of transfer amounts, the intended value was determined by the Cost of the Diet analysis undertaken by the Federal Ministry of Health/Ethiopian Public Health Institute (EPHI) with technical support from WFP and was based on household size: eligible households with up to two members received a value of USD 14, households with three to five members received USD 17, and households with six and more members received USD 21. Compared to the average values received during the programme period, it becomes apparent that voucher sizes did not align with the household sizes reported at end-line, especially for smaller families. Thus, the

voucher amounts seem to have been set based on the household size at enrolment and not been updated according to any additional births during the beneficiary period. This may be another reason why some beneficiaries considered the voucher inadequate to cover their family's fresh food consumption needs.

Further evaluation questions covered the availability of, access to and demand for fresh foods. The programme was effective in stimulating the supply side to meet the specific nutrition needs of the beneficiaries. Fresh foods in form of dark leafy greens, orange/dark yellow fruits and vegetables, as well as animal sourced products were largely offered by FFV traders. While seasonality was less considered a problem in the supply, it added to the issue of price fluctuations for key products. High prices were also partly ascribed to trader behaviour, e.g., through potential collusion or unfair price setting. Access to markets was mixed for beneficiaries: A considerable portion of beneficiaries faces long walking distances to markets, which implies both restrictions to access as well as safety concerns to female beneficiaries. Lack of transportation has been raised not only as a restriction during COVID-19, but also in general. In line with the perception of relevance of the vouchers, demand for fresh foods purchased with the same was high. Main consumed food groups of PLW and children in beneficiary households include grains, roots and tubers, pulses and seeds, vitamin A rich foods and other fruits and vegetables. There also seems to be a high level of knowledge, as well as reported practices of favourable child feeding, such as exclusive and continued breastfeeding. This implies that SBCC messages were understood and followed widely by caregivers. One needs to keep in mind, though, that reported practices could not be verified because of the remote data collection due to COVID-19.

In terms of dietary outcomes, we investigated the minimum acceptable diet through minimum dietary diversity and meal frequency for children as well as dietary diversity of PLW. When comparing to the targets set by the programme, we found that child outcomes were mostly met, while women's dietary diversity targets were largely missed. These results may have likely been affected by the pandemic, such that clear conclusions cannot be reached. However, these results point towards a potential buffering effect of mother's diets compared to those of their children. Further, we find that with awareness as well as attendance of SBCC activities, women's dietary diversity improves. This highlights the importance of awareness-raising and behavioural change not only for child feeding, but also for women's diets.

Sustainability: There is a high willingness to continue the programme from the governmental, as well as donor side, with concrete negotiations ongoing since the end of the programme. Government capacities have been fostered in skills including nutrient gap cost and non-affordability of nutritious diet analysis to inform the re- adjustment of the FFV transfer value, and the contents of the SBCC implementation.

In terms of potential for sustained behavioural change after programme end, this evaluation could only give indications about the eating behaviours shortly after leaving the programme (while the programme itself was still ongoing). We find that while there seems to be high understanding of the importance of dietary diversity and willingness to continue consumption of fresh foods, financial constraints are a prohibiting factor. In terms of dietary diversity outcomes, there seems to be a trade-off between women's and children's diet: It appears as if mothers prefer to provide fresh foods to their children, even when they are older, instead of consuming a diverse diet themselves. While children's diet is enriched, mothers' diet is less so.

Recommendations to be considered for a possible continuation or expansion of the FFV programme

Recommendation 1: WFP should strengthen technical IT capacities for smooth registration, transfer and redemption process. Adaptations to reliability of mobile phone network to be considered.

Recommendation 2: WFP should regularly assess the adequacy of transfer value in light of price fluctuations, and possibly of the beneficiary households' affordability.

Recommendation 3: WFP should consider conducting an impact evaluation with control group design in order to rigorously assess causal attribution of the, so far, promising results of the pilot programme.

Recommendation 4: Trade authorities, with WFP support, should strengthen accountability mechanisms for traders at the local markets.

Recommendation 5: WFP should focus on awareness-raising activities on nutrition for PLW and consider gradual phasing out.

Recommendation 6: WFP should investigate interactions with other WFP and government programmes (such as the Productive Safety Net Programme transfers and home gardening programmes).

Recommendation 7: Given the importance of gender-sensitive programming, WFP should continue to investigate interactions and possible synergies with other governmental and private sector actors in order to reinforce the Digital and Financial Inclusion of Women through phone-based voucher programmes.

1. Introduction

1. This Evaluation Report is for the end-line activity evaluation of the **Support for Strengthening Resilience of Vulnerable Groups in Ethiopia: The Fresh Food Voucher (FFV) Programme Expansion in Amhara Region**. The FFV Programme Expansion in Amhara Region resulted from the successful pilot launched in three woredas (Habru, Raya Kobo, and Dessie Zuria) in 2017. Following the interest of the Government of Ethiopia (GoE) to further expand the program, and the support of the German Federal Ministry for Economic Cooperation[1] (BMZ) and the German Development Bank[2] (KfW), the FFV Expansion Programme was launched in January 2018 expanding to an additional four woredas (Dawa Chefa, Kalu, Seqota and Mekdela), but first disbursements took place in the second quarter of 2018. The expansion phase ended in December 2020.
2. The evaluation is commissioned by World Food Programme (WFP) Ethiopia Country Office (ETHCO) and covers the period from January 2018 to December 2020.
3. This end-line evaluation follows the two objectives: accountability and learning. It shall inform on the performance and results of the FFV Programme Expansion in Amhara Region and help explain why certain results occurred and advise on good practices and lessons for future programming. It needs to be pointed out, however, that the evaluation is severely limited by the COVID-19 pandemic because of its effects on outcomes and impacts observed through primary data collection which might have been very different without the pandemic. In addition, the COVID-19 situation also imposed restrictions on the feasibility of data collection tools. In particular, only short phone interviews were permitted, and no face-to-face household surveys were allowed, which limits the richness and level of detail of the collected information. It also implied that important outcomes such as anthropometric measurements of children could not be collected. Since the expansion phase ended in December 2020 a delay in data collection could not be considered. Nevertheless, the evaluation can still provide insights about lessons learnt in the design and implementation of an innovative voucher programme, including the readiness to react to an external shock such as the COVID-19 pandemic.
4. The evaluation covers questions pertaining to the relevance, effectiveness, efficiency and sustainability of the programme.
5. The expected users for this Endline Report are WFP ETHCO, its partners in primis, Seqota Declaration, Productive Safety Net Programme (PSNP) V and National Nutrition Programme (NNP) II, as well as the East African Regional Bureau, other WFP Country Offices in the Region, the Office of Evaluation of WFP and WFP Headquarters. WFP and government stakeholders are expected to use this evaluation to help determine the potential continuation and expansion of the FFV programme. BMZ/ KfW as a main donor for the expansion of the programme is expected to use the evaluation findings to derive lessons learned and inform future funding.

1.1. Overview of the Evaluation Subject

6. The subject of this end-line activity evaluation is the FFV Programme Expansion in Amhara Region from January 2018 to December 2020. Embedded in the PSNP and the Seqota Declaration under the NNP[3], key partners are the GoE, Ministry of Agriculture and Natural Resources, Ministry of Livestock and Fishery Resource Development, Ministry of Health, Ministry of Water, Irrigation and

[1] Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ)
[2] Kreditanstalt für Wiederaufbau (KfW)
[3] These entities are further discussed below under 2.1.

Electricity, Ministry of Education, and Ministry of Labour and Social Affairs. The expansion is funded by a 7 million Euro grant from BMZ/KfW[4]. WFP was the main implementing organization of the FFV with HelloCash as the service provider and woreda-level agencies and Health Extension Workers (HEW) engaged to implement the programme on the ground.

7. A large component of the FFV consisted of government and private partnerships with national financial service providers, trade authorities at district level, and ultimately with small retailers/suppliers. In particular, WFP engaged and trained HEW and development agents (DA) for the implementation of the Social and Behaviour Change Communication (SBCC) component. HelloCash by Lion International Bank and Belcash is the private mobile money wallet utilized by the FFV programme. Further, fresh food traders were engaged as participants of the programme.
8. The FFV Programme Expansion covers the original three pilot woredas (Habru, Raya Kobo, and Dessie Zuria), as well as four additional woredas in Amhara: Dawa Chefa, Kalu, Seqota and Mekdela (see Figure A 1). As described in section 1.2., especially the Eastern Amhara Region is particularly affected by food insecurity.
9. The main beneficiaries of the programme are households with pregnant and lactating women (PLW), and with children aged 6 to 23 months[5]. Given the design of the intervention, namely vouchers transferred digitally to mobile phones using SMS messages, an important eligibility criterion is the ownership of a mobile phone. While mobile phone ownership is generally linked to wealth, an assessment of coverage during the first phase of the pilot revealed adequate access to phones for beneficiaries. Therefore, this was not part of this evaluation.
10. Through the vouchers, the programme aimed to complement the need for diet diversity of PSNP beneficiaries, as it was designed as a top-up to cover nutritious needs. The total number of targeted households was 27,000 by the end of the programme, with 9,000 from the pilot woredas and 18,000 from the expansion woredas[6]. In August 2020, 96% of the target has been reached (see Table A 1 for details).[7]
11. The FFV programme's main objectives are, in PSNP woredas, to 1) increase the dietary diversity of PLW and children under two years in households eligible for PSNP programming, 2) stimulate the local fresh food market (i.e., fruits, vegetables and eggs) and 3) strengthen social and behavioural change to generate more demand for fresh food in seven woredas in the Amhara Region
12. To reach these objectives, the programme includes a three-pronged approach to availability, access, demand and utilization of fresh foods: the main innovation of the programme are mobile money vouchers to beneficiaries redeemable for vegetables, fruits and animal source foods to enhance access to fresh foods. This component is coupled with retailer engagement and trainings for fresh food suppliers/retailers in rural markets, as well as SBCC activities. Combined with community mobilization and partnerships with governments as drivers of change, this shall lead to PLW and children under 2 adopting a healthier, more diverse diet (see Annex 4 for Theory of Change). The ToC is soundly based on the available evidence from similar interventions. In particular, evaluations of the PSNP recommended integrating gender sensitive SBCC and integrating HEW in the programming. Another meta-evaluation of fresh food voucher programs in several countries found that e-vouchers reduce administrative costs and ensure timely delivery.[8]

[4] A breakdown of the costs was not made available to the evaluation team.
[5] Children under 6 months are expected to be targeted through breastfeeding women.
[6] Numbers taken from WFP Annual Report to KfW 2019 (version 1.0 from 20.01.2020).
[7] From WFP Progress Report to BMZ/KfW for the period January - August 2020.
[8] See a list of reviewed evaluations in Annex 3.

13. The evaluation builds upon the ToC in the formulation of evaluation questions and respective quantitative and qualitative tool. In particular, implementation fidelity and changes to the original design will be investigated. Further, due to unintended disruptions, e.g. due to the COVID-19 outbreak and containment measures, some underlying assumptions may be affected. For example, supply of products including fresh foods may be hampered by lockdowns; income flows of households may be disrupted in the short term and intra-household food allocation adjusted as a consequence.
14. The voucher value was designed to fill the gap between the affordability and the cost of nutritious diet of the entire household, and therefore was based on the household size, while accounting for the PSNP transfer (assuming that the latter was available): FFV-eligible households with up to two members received a value of USD 14, households with three to five members received USD 17, and households with six and more members received USD 21. Households receive these transfers on a monthly basis to their mobile money wallet, provided by HelloCash. The voucher is received on monthly basis via SMS messaging, can be redeemed, partially (in different times during the months) or totally (in all its amount at once), at retailers with an accepted trader license at the rural market.
15. SBCC activities included the distribution of information material through HEW and in communities, coffee conversations and community theatres, as well as radio broadcasts of some of these activities, training of HEW, and cooking demonstrations conducted by HEW. Due to the COVID-19 outbreak, coffee conversations and cooking demonstrations were suspended in February 2020. To counter this loss, new SBCC activities in form of SMS and audio messages on mobile phones were rolled out in all woredas from July 2020 onwards.
16. Retailer engagement and trainings were conducted to address supply chain inefficiencies and to improve availability, prices, and quality of fresh foods and services to vulnerable households. Food traders at selected markets were trained in adequate food handling, storage and transportation.
17. The programme targets women as the main beneficiaries by improving access to nutritious foods for PLW and children caregivers (e.g. their mothers) as they obtain access to a mobile money account. However, the programme did not foresee any monitoring of who is using the mobile money wallet and redeeming the voucher at the market, for example. Rather, the focus of voucher usage was on household level.

1.2. Context

18. According to the most recent State of Food Security report (FAO, IFAD, UNICEF, WFP and WHO, 2020), more than 63 million people in Ethiopia are food insecure and 15.4 million are severely food insecure. With 40% of children being stunted across the country, the associated cost is an equivalent of 16.5% of the GDP yearly (EPHI and WFP, 2013).
19. Particularly in the eastern part of the Amhara Region, food insecurity remains a chronic issue. Compared to the national averages, Amhara Region fares slightly worse in the prevalence of malnourished children below 5, including stunting (41%). (UNICEF, 2019) A report by IFPRI in Amhara in 2019 finds that merely 3% of children between 6 and 23 months have a sufficiently diverse diet, while even less of women meet dietary diversity requirements. Animal-sourced foods such as dairy, meat and eggs, as well as vitamin A rich fruits and vegetables, are rarely consumed among women and children (Hirvonen K, 2019).
20. Ethiopia faces deep-rooted obstacles to gender equality and the empowerment of women (GEEW) within society. A study on economic empowerment of rural women in Amhara, shows that only

21% of rural women in the region are empowered, according to the Women Empowerment in Agriculture index[9]. Gaps in equality persist in Amhara Region also in terms of health-related knowledge. Further, Amhara ranks among the lowest rates in the country on economic empowerment, measured by consistent paid work. In 2016, 21% of women in Amhara owned a mobile phone compared to 48% of men (MOWCY, UNICEF Ethiopia and SPRI, 2019).

21. In an effort to counter issues of food security, poverty, vulnerability and to provide a more efficient method of distributing emergency food aid, the GoE developed its flagship PSNP. The programme has various objectives, aimed at improving livelihoods, food security, nutrition and resilience to shocks of rural vulnerable households. WFP supports the PSNP programme with multiple complementing mechanisms including the provision of in-kind assistance, and the FFV programme (WFP, 2020). Additionally, the PSNP has strong linkages to the NNP, notably the Seqota Declaration, a high-level commitment which has the specific goal to end child undernutrition by 2030 (Federal Republic of Ethiopia, 2016). Apart from the FFV, WFP engages in Amhara Region through conditional in-kind food or cash-based transfers to PSNP beneficiaries and livelihood support for refugees and host populations.
22. As part of the "Fill the Nutrient Gap" Analysis in 2020, the Ethiopian Public Health Institute (EPHI) together with WFP assessed the issues around availability, access, affordability and intake of nutritious foods. It found that 3 out of 4 households could not afford a diet that meets all nutrient requirements, with breastfeeding women and their children among those with the highest risk of an inadequate diet. Further, while households mainly purchase fresh foods such as fruits, vegetables and meat at the market, the prices of these goods have increased in recent years (EPHI and WFP, 2020).
23. Other key actors in the region include the United States Agency for International Development (USAID), which is implementing the Growth through Nutrition (2016-2021) multisectoral project, together with several partners.[10] Other major nutrition and food security interventions are implemented by UNICEF, the World Bank, the Global Alliance for Nutrition, and Scaling up Nutrition.
24. Following the first confirmed COVID-19 case in Ethiopia in March 2020, the government was quick to close schools, and suspend public events and gatherings. Regional governments followed suit and curbed the internal movement of individuals, through travel bans, lockdowns and suspension of public transportation services.[11] Besides business closures and the slowdown of economic activity, the suspension of classes also disrupted school feeding programmes. The distribution of food assistance became severely constrained because of the restrictions on the movement of people, bans on large public gatherings, and scarce personal protective equipment to deliver goods safely (UNICEF , 2020). Due to the 'triple menace' of the pandemic, heavy flooding, and the desert locust infestation, food prices in Ethiopia rose considerably in the first months of 2020[12].
25. The recent conflict in Tigray Region has had ripple effects on the neighbouring regions of Afar and Amhara, displacing several thousands of people. The programme woredas were also partially affected by the conflict, as regional authorities have reported the return of 900 internally displaced

[9] The Women Empowerment in Agriculture index measures empowerment across five dimensions: production, resources, income, leadership, and time.

[10] The partners include Save the Children, World Vision, Jhpiego, Population Services International, Land O' Lakes International Development, Tufts University, The Manoff Group, Ethiopian Orthodox Tewahedo Church-Development Association, Fayyaa Integrated Development Organization, and The Ethiopian Muslims' Relief and Development Association.

[11] Ezega News. (2020) 'Ethiopian Regional States Impose Travel Ban to Halt Spread of COVID-19' March. 30th, accessed 22/3/2021

[12] https://www.wfp.org/news/food-rations-cut-refugees-eastern-africa-coronavirus-stretches-resources, accessed 22/3/2021

persons (IDP) to their place of origin in Raya Kobo. In Seqota, all communication channels were disrupted, hampering project activities and impeding remote data collection for the evaluation in that region. (OCHA, 2020)

1.3. Evaluation Methodology and Limitations

26. As per the Terms of Reference, we aimed to evaluate the programme through the lens of the Development Assistance Committee (DAC) criteria of Relevance, Effectiveness, Efficiency, and Sustainability.[13] Among these, we consider Effectiveness and Sustainability as the main criteria of interest for this evaluation. A mixed-methods approach, using both quantitative and qualitative analyses from primary and secondary sources is employed. A detailed overview of the evaluation criteria, main and sub-evaluation questions, methods and tools used is given in the evaluation matrix in Annex 5. The full methodology is presented in Annex 6.
27. The quantitative tools include primary data sources such as a survey with beneficiary households and a survey among participating traders. Secondary data sources include the HelloCash analytics, the SCOPE beneficiary database and programme enrolment lists updated as of October and December 2020. Quantitative tools used to collect primary data are surveys with (i) beneficiary households and (ii) participating FFV traders. Qualitative methods used include key informant interviews (KII) with implementers and partners, and in-depth interviews (IDI) with beneficiaries and traders. All quantitative and qualitative data was collected remotely via phone.
28. Evaluation questions on relevance are covered through qualitative analysis. To assess the effectiveness and efficiency of the programme, we use both quantitative and qualitative tools and triangulate the findings. We cross-checked certain information collected from household interviews with data from M&E systems. The qualitative data from IDIs give more in-depth insights into experiences with the FFV programme, complementing the aggregated results on usage of the vouchers and the demand and supply of fresh foods from the survey data. Evaluation questions on sustainability are primarily answered through qualitative methods; the sustainability in terms of nutrition outcomes of households no longer benefitting from the programme are assessed through survey results.
29. Given the gender focus of the programme, GEEW aspects are mainstreamed into different evaluation criteria, particularly relevance and effectiveness. Both gender dimensions of beneficiaries and children are covered. Particularly, the focus is on women's decision-making power and gender equality in nutrition, i.e., no preferential treatment of male family members in relation to diets.
30. We were not able to design an impact evaluation in order to explore causal effects as there is no adequate information on a credible comparison population currently available.[14,15] (In the absence of causal impact estimates, also no cost-benefit analysis is possible.) Instead, the methodology for

[13] For more details, see http://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm.

[14] WFP was able to retrieve lists of PSNP beneficiaries from neighbouring woredas where the programme was not rolled out. However, the key data to determine a credible comparable control group based on eligibility criteria as of year 2018 or 2019 are missing. The database contains information as of today, but not about the population that would have been eligible in 2018 or 2019.

[15] We also explored a possible before-after comparison design for key outcomes of beneficiaries in order to shed light on the performance of beneficiaries with respect to nutrition outcomes targeted by the implementation. WFP has been conducting annual PDM surveys with small samples of beneficiaries since 2019 in a subset of woredas, including information on food-security and nutrition of mothers and their children. However, the compiled PDM databases were not complete and therefore not usable for sampling purposes.

the quantitative analysis of nutrition outcomes contains of descriptive statistics for the full sample and by various subgroups. Additionally, we employ regression analysis to control for household and individual characteristics and explore associations between the outcomes and potentially relevant indicators.

31. For the beneficiary group, a representative sample size of 1380 was drawn from the entire beneficiary population across the programme regions Habru, Raya Kobo, Dessie Zuria, Dawa-Cheffa, Mekdela and Kalu[16,17]. As the main beneficiaries, the sample consists of female respondents in households with children up to the age of four years. As for the trader's survey, we interviewed 140, i.e. more than 72% of the HelloCash registered traders in the programme across the same six woredas.
32. The 14 respondents for beneficiary IDI were selected from the database of beneficiaries not participating in the household survey in order to avoid that the same person was interviewed twice, since both data collections happened in parallel. (The respondents were randomly selected from the pool of beneficiaries not selected for the phone interviews.) The same strategy was followed to select the 14 trader respondents for IDI. KII were conducted with high-level representatives of the programme and local stakeholders including programme coordinators, as well as nutrition, agriculture, and trade officials on woreda level.
33. Quantitative and qualitative data collection took place entirely remotely, i.e. by phone. A mix of male and female enumerators located in Addis Ababa conducted phone interviews in Amharic or Oromiffa language. The enumerators were trained on interview conduct via phone and ethical concerns, the content of the survey and use of the CAPI programme. Survey fieldwork was conducted between 10 and 14 December 2020, qualitative interviews took place 31 December 2020 to 19 January 2021.
34. The full Quality Assurance (QA) plan was detailed in the inception report. The Center for Evaluation and Development (C4ED) carries out the evaluation in compliance with the WFP's Decentralized Evaluation QA System to ensure the credibility and utility of evaluations and facilitate learning. The QA covers various inter-related dimensions, including the quality of deliverables, the organization of the process, the utility of the evaluation and the timeliness and management of relations with relevant stakeholders.
35. The mode of data collection (telephonic interviews) severely hampers the reliability of the data (see below and Annex 6 for limitations). Nevertheless, C4ED ensured reliability as far as possible through a series of high-frequency data checks during data collection. Further, the triangulation of different primary and secondary, quantitative and qualitative data sources was employed to increase the validity of the results.
36. During the inception phase, a number of ethical issues and related safeguarding measures have been considered for the preparation/design, data collection, data analysis, reporting and dissemination. These are summarized in Table A13 of Annex 7. These issues were monitored and managed during the implementation of the evaluation.

Limitations

[16] The seventh woreda, i.e. Seqota, was dropped from the sample because of the ongoing outage of the phone network connectivity in the area due to a conflict in the neighbouring Tigray Region.
[17] A detailed description of the sampling and stratification procedure can be found in Annex 6.

37. COVID-19 imposed serious limitations on the data collection process. Given the COVID-19 pandemic situation in Ethiopia, WFP had decided that face-to-face interviews were not feasible. Instead, interviews were conducted by phone. This was possible because the FFV programme uses phone numbers to provide the monthly transfers, via SMS-messaging. Phone interviews come with several crucial limits to interview length (questionnaires were cut to a length of maximum 30 minutes), scope (anthropometric data could not be measured), and reliability (e.g. respondents' understanding of questions, observation of body language, non-verbal cues or the use of visual aids, lack of observation of surroundings, i.e., other present family members and their impact on the respondent's ability to speak frankly).
38. Another issue pertaining to data collection via phone is the issue of reachability of respondents, both in terms of obtaining the correct phone numbers and connectivity. In order to tackle these problems, we liaised with woreda officials and HEW before the start of the data collection for the verification of phone numbers and reaching out to respondents beforehand to ensure awareness of the upcoming survey.
39. The conflict in the Region of Tigray ongoing at the time of the evaluation generated spill-over effects in neighbouring regions and raised challenges to the data collection and analysis. Due to a shutdown of the phone network in Seqota woreda right before the start of data collection, conducting phone surveys in Seqota woreda was deemed impossible[18]. Similarly, to the consequences of the COVID-19 pandemic, the Tigray conflict can have significant effects on access to fresh foods and diet-related practices in programme areas.
40. Given the focus of the programme on women and children, data was only collected from female respondents. A limitation in this regard is that no information on the male perspective or intra-household decision-making, as well as perceptions of child feeding could be collected. Similarly, any perception regarding gender empowerment (e.g. through financial inclusion) is limited to the view of female respondents.
41. Further, given the limited possibilities pertaining to the phone-based data collection, it was not possible to collect data from non-FFV participants.

[18] After an assessment of the situation, WFP ETHCO and RBN decided on 2 December 2020 that, given the tight timeline for data collection until the end of 2020, this woreda had to be dropped from the sample.

2. Evaluation Findings

42. The evaluation findings and the evidence to substantiate them are presented below. They are structured as a response to each evaluation question in turn.

2.1. Evaluation Question 1.1: Is the project aligned with the national government's policies and strategies to support the reduction of stunting via increased diet diversity?

43. As part of its national development agenda, Ethiopia has been implementing different strategies and programs to ensure food and nutrition security, such as the Cost of Hunger Study, the Food Security Strategy, National Nutrition Strategy (NNS), NNP, the Seqota Declaration, Nutrition Sensitive Agriculture Strategy, School Health and Nutrition Strategy and the PSNP, through multi-sectoral nutrition coordination and integration. The government also incorporated nutrition, with particular attention to the reduction of stunting, into its 5-year Growth and Transformation Plan.
44. The GoE has continued its commitment to nutrition by developing the 2nd phase of NNP (NNP II, 2016-2020) that, by adopting a life-cycle approach to address malnutrition, gives great emphasis on the crucial periods of pregnancy and the first two years of a child's life. Strategic objectives include i) improving nutritional status of women of reproductive age and PLWs by promoting maternal nutrition and related SBCC, adequate intake of diversified foods, personal hygiene and support to women's empowerment; ii) improving nutritional status of infant and young children 0-23 months by promoting optimal breastfeeding practices for infants 0-6 months, complementary feeding for children 6-23 months, and key actions for diversification and utilization of complementary foods at household level. The NNP II aims to reduce stunting prevalence among under-five children from 40% to 26% and increase the proportion of children 6-23 months with minimum dietary diversity score from 5% to 40% by 2020. (Federal Republic of Ethiopia, 2016)
45. A component of NNP II is the high-level commitment called The Seqota Declaration of 2015, which aims to end stunting in children under two years by 2030. Innovation is central to achieving this goal and requires nutrition smart interventions across multiple sectors. Strategic objectives include i) improving the health and nutritional status of adolescents, women and children under two years, ii) ensuring 100% access to adequate food all year around, focusing on production and consumption of fruits and vegetables, staple crops and pulses, meat and dairy products, iii) ensuring zero post-harvest food loss, focusing on proper postharvest handling, storage, processing and marketing of agricultural products and developing quality assurance and food safety guidelines, and iv) promoting implementation of gender-sensitive social safety net programs. Many of these interventions are driven by SBCC strategies. (Federal Republic of Ethiopia, 2016)
46. In order to provide the necessary legal and institutional framework for national nutrition planning, implementation, and coordination in the country, the GoE launched in 2018 its first Food and Nutrition Policy (FNP). The policy covers key dimensions of food and nutrition security including availability and accessibility of adequate food, consumption and utilization of a diversified and nutritious diet, safety and quality of food throughout the value chain,

postharvest management, and food and nutrition literacy to all Ethiopians. (Federal Republic of Ethiopia, 2018)

47. The FFV programme in the PSNP districts combines several aspects of the above mentioned key governmental policies and strategies. The programme's relevance for stunting reduction has been established with key government stakeholders. By focusing specifically on PLWs and children six to 23 months, the project recognizes the critical importance given to the first 1,000 days of child's life. Its primary objective is that of improving dietary diversity among vulnerable households with PLW, and children under two, in rural woredas covered by PSNP, and secondly, stimulating the local fresh food market. As a price subsidy scheme, the FFV aims to ensure financial accessibility to fruits, vegetables and animal sourced food, topping up the PSNP transfer which is limited to grains, oils, and sugar. Ultimately, FFV value contributes to fill the gap between the limited affordability and the cost of a nutritious diet in the PSNP woreda eligible household. Through the integrated SBCC component, the financial access is supported by awareness and knowledge creation of PLW, while the availability of nutritious food is targeted though improvements of the supply capacity by trader engagement and training.

Key findings and conclusions – Question 1.1

- Key national policies and strategies include the NNS, NNP, Seqota Declaration, and the PSNP.
- These policies acknowledge the importance of nutrition interventions for PLW and children under the age of two for stunting reduction in Ethiopia.
- The FFV programme follows these guidelines with an integrated and PSNP-complementary intervention including the access to nutritious fresh foods, demand creation and awareness raising with PLW and the care-givers of children under two years of age, and strengthening the provision of nutrient-dense foods at the local market.

2.2. Evaluation Question 1.2: How relevant has the approach been to beneficiaries?

48. We used qualitative IDI with beneficiaries to gauge the relevance of the programme from their perspective across four themes: (i) the alignment of FFV recipient and decision-making regarding food purchases in the household; (ii) the usefulness of the vouchers and information received via SBCC; (iii) its usefulness in light of the COVID-19 pandemic; and (iv) the relevance for male versus female children.
49. Many beneficiaries stated that they had initially registered with a family member's or neighbour's phone, suggesting that the vouchers were received by someone else in, or in some cases, outside of the own household. Nevertheless, qualitative data suggests that the vouchers have been accessible to the PLW for own use and for their children. Respondents mentioned that they are involved or mainly responsible in the decision of voucher utilization. Thus, voucher receipt reception and utilization seemed to have been closely aligned. It should be pointed out, however, that a closer tracking of the relation between direct recipient and intended beneficiary could have been valuable to ensure that the flow in transfers was not interrupted, e.g. because of the phone holder migrating or being otherwise unattainable.
50. Beneficiary IDI revealed a sense of usefulness of the programme, particularly the voucher component. Improved food intake, access to healthy foods, stimulation of agricultural production, economic gain, women empowerment, getting introduced to technology and

contribution to climate change were the benefits of the FFV programme listed by the respondents during the interviews. They further welcomed the transfer modality, i.e. vouchers redeemable on certain goods as opposed to cash.

51. Beneficiaries demonstrated an understanding of the benefits from consuming fresh foods and favoured the FFV for granting financial access to healthy foods. One beneficiary interview reads:

"Before the programme started, I yearned to buy these foods, but I could not buy them (…) due to shortage of money; in addition, when I want to fulfil other consumption products such as coffee, red pepper, salt and the like, it is very difficult to buy vegetables. (…) But since I have got a support from the government for the vegetables (…) I use the money that I got from the government."

(Beneficiary, female, Dessie Zuria)

52. Generally, beneficiaries deemed the digital voucher system clear and simple to use after the training received at the beginning. Further support throughout the transfer period allowed beneficiaries to comfortably navigate the vouchers. An additional advantage during the COVID-19 pandemic evolved through the cashless payment method which allowed to reduce physical contact. No clear results emerged if the programme otherwise aided households in accessing fresh foods during the pandemic.
53. A predominant outcome from IDI was that the number of transfers was perceived inadequate for the fresh food needs due to large household sizes, but also due to frequent price fluctuations. Some impressions of higher prices at FFV traders compared to the market price level were also voiced. What remains unclear from the IDI is if those households perceiving the voucher amount as inadequate are non-PSNP beneficiaries. As a top-up to the PSNP, the amount of the FFV may not be sufficient to cover nutritious needs if the household is not receiving a PSNP transfer. As will be discussed below in section 2.4., only a low number of FFV households also received PSNP support. Therefore, this explanation may stand to reason in cases where the voucher amount was considered too low.
54. Some beneficiaries suggested that an expansion of possible products such as ginger, garlic, oil and wheat would help them cover their nutritious needs. Others claimed that important products, such as apples and mangos, are not provided at their FFV traders.
55. Another barrier to the use of FFVs mentioned was the initial misconception attributed to HelloCash financial services, perceived as a religion-linked organization in some areas. This led some churches to ban the involvement of people in the FFV programme and induced some beneficiaries to hide their participation to the program because of fear to be discriminated. While this has been mitigated through awareness-raising within communities and specifically beneficiaries, deep-rooted misconceptions may remain with religious or other community leaders and may serve as a cautionary tale for other woredas.
56. The predominant view of beneficiaries was that fresh foods are mainly important for children where women and other adults in the household consume after the children get the adequate amount. With regards to the gender of the child, no differences were reported in the type of fresh food served to male or female children. One beneficiary explained that feeding patterns for male children used to differ, but that younger generations do not follow these rules any longer:

"Yes, we believe in our culture that boys would be full-hearted/brave and with good moral if they drink milk. (…) It doesn't mean she (the girl) doesn't have to drink at all. It means that she drinks

less milk compared to boys. (…) Today, we make it equal. (…) I used to listen my mother and grandmother saying milk and meat are good for boys (…). If I gave birth to a baby girl, I would not treat her differently." (Beneficiary, female, Dessie-Zuria)

57. On SBCC activities, reviews were mixed as different woredas experienced roll-out and interruptions differently. Some beneficiaries stated that they have not seen any food preparation demonstrations or other activities, and instead received arguably less helpful oral advice on how to prepare foods.

Key findings and conclusions – Question 1.2

- Household decision making regarding food purchases seems to be largely aligned with handling the mobile money vouchers. Follow-up on who receives and handles the vouchers would be valuable to closer track this association.
- The programme is useful to support consumption of fresh foods. Main barriers include the voucher amount and community pre-conceptions about the service provider. With regard to the voucher amount, a likely explanation is that those households have not benefitted from the PSNP, such that the FFV alone was not sufficient to cover nutritious needs.
- Within the household, consumption of fruits and vegetables is first granted to children and subsequently to women and other adults. There seem to be no differences in feeding practices towards male and female children.
- While messages received via SBCC seemed to be informative and applicable for beneficiaries in general, experiences differ due to differential roll-out and suspension of interactive programmes during the pandemic.

2.3. Evaluation Question 1.3: How did the programme interventions adapt to the COVID-19 pandemic?

58. The pandemic outbreak brought about inevitable changes in the programme implementation. Due to COVID-19 restrictions, the initial plan of conducting assessment and registration of beneficiaries once every three months was disrupted, as well as the house-to-house follow-up and assistance that used to be given to beneficiaries regarding voucher usage and types of fresh food items.
59. Notably, cooking demonstrations and other gatherings were suspended in February 2020. To counter this loss, new SBCC activities in form of SMS and audio messages on mobile phones were developed and rolled out in all woredas from July 2020 onwards. Further, mass media channels, such as radio and television, were engaged in December 2020 to disseminate SBCC messages.
60. In collaboration with the Trade and Industry Coordination Office, market days were added to minimize the crowdedness of marketplaces. Face mask and hand sanitisers were also made available by WFP for times of interaction in the market.
61. Given the unpredictability of the COVID-19 situation, the programme was able to quickly redistribute resources to continue the intervention and adapt some programme components in 2020.

Key findings and conclusions – Question 1.3

- The frequent face-to-face interactions during registration, market assessments, and follow-up visits were reduced since the COVID-19 outbreak.

- Similarly, interactive SBCC activities were suspended in February 2020. New concepts were developed and starting from July 2020, SMS and audio messages on mobile phones were disseminated. Expansion of SBCC messages through mass media outlets followed in late 2020.

2.4. Evaluation Question 2.1: Did the project reach the intended beneficiaries with the right mix of assistance?

62. To answer this evaluation question, we first uncovered if those enrolled in the program fulfil the eligibility criteria, i.e., are households with PLW or children of 6 to 23 months. Further, as the FFV was intended to be disbursed as a top-up to PSNP support, we observed if the household received a PSNP transfer as an additional targeting factor. From our sample of 1380 households, 794 (57.5%) were still enrolled in the programme at the time of end-line (December 2020), while 586 (42.5%) were no longer enrolled (graduated). Thus, we examined eligibility criteria for both groups.
63. As Table 1 indicates, 97.6% of those enrolled in December 2020 exhibit eligibility characteristics. Notably, only around 20% of the sampled beneficiaries received a PSNP transfer in the last year, indicating that this criterion was not closely followed up on at enrolment. However, as the programme intended to target both PSNP and non-PSNP households with PLW or children under 2, this result rather indicates that PSNP coverage may have been lower than expected. Nevertheless, it highlights the fact that for the majority of the beneficiary households, the FFV did not function as a top-up to the PSNP transfer.
64. Further, almost 75% of households that are no longer enrolled in the programme have a pregnant or lactating woman or children below 2 years, implying that they would still be eligible. Of all graduated households, almost 70% have a pregnant or lactating woman and around 35% have children between 6 and 23 months. An apparent consideration may be that those households may have a child that has recently surpassed the eligible age but have become eligible again due to a new pregnancy. To check this, we look at graduated households that are still eligible and have a child above 23 months. Indeed, 62.4% of households have an older child. Still, for the remaining 37.6% it appears that enrolment has been terminated pre-maturely. While the majority (80%) of those was registered in 2018, there has been no indication in KII with programme officials or otherwise that enrolment ended due to other time restrictions. Thus, no clear conclusion can be reached regarding the reasons for this large share of eligible but graduated households.

Table 1: Eligibility criteria of enrolled and graduated households

Eligibility Criteria	**% of households**	
	enrolled	**graduated**
Pregnant or lactating women	91.9	69.3
Children 6-23 months	87.3	35.8
PLWs + children 6-23 months	97	74.7
Transfer from PSNP	19.4	19.8
Total number of households	**794**	**586**

65. The next step was to investigate if the households received the FFV transfer as intended, i.e. a monthly value based on the household size. Table 2 cross-checks statistics from different

sources: Over 95% of (enrolled) respondents indicate that they received a FFV transfer in the past 6 months. HelloCash Analytics for beneficiaries were available from June 2018 up to September 2020 (i.e., 2 months before the end-line survey took place). Observing the six-month interval April to September 2020, around 58% of the sample received a transfer regularly each month. The average number of transfers received during this period was 4.4.

Table 2: Received transfers

FF voucher received	Number of households	% of households
In the last 6 months	757	95.5
Every month (April 2020-Sept 2020)[19]	598	58.6

66. The transfer value was determined by the Cost of the Diet analysis undertaken by the Ministry of Health/EPHI with technical support from WFP and accounts for the food prices and the cost of a nutritious diet in a household living in the PSNP area, while accounting for the nutrition requirements of a pregnant/lactating woman, an adult man having intensive physical activity, an adolescent girl, a school child, and a child with less than two years of age (Bose I., 2019).
67. Assessing the correct transfer value proved challenging for several reasons. First, the transfer value, which is reported in Birr in the HelloCash database, was regularly adjusted based on the exchange rate between USD and Birr. Thus, the value may have fluctuated throughout the reported period. During qualitative interviews, it was also mentioned that the transfer values may be adjusted based on regular price assessments on the market. Further, it is unclear if and when the disbursement is updated based on updated household size information, in case the household size increases after registration. Therefore, the exact expected transfer value based on the household size at end-line could not be defined. Nevertheless, we present statistics of the average Birr value received between June 2018 and September 2020. Table 3 shows the average values based on household sizes at end-line and at enrolment in original Birr as reported in the HelloCash Analytics, as well as USD (based on an average exchange rate). Per voucher, households with up to 2 members (at end-line) received an average amount of 516.49 Birr (16.92 USD), households with 3 to 5 members 489.05 Birr (16.02 USD), and households with 6 or more members an average amount of 565.46 Birr (18.52 USD). The average amount drops from small to medium sized households, which may indicate that additional household members (e.g., when a child was born) were not readily accounted for in the disbursement of vouchers. Indeed, when comparing to the household size at enrolment, the intended pattern of increasing voucher amounts by household size emerges. This could be another indication why transfer amounts were considered inadequate by some beneficiaries.

Table 3: Average Birr and USD value received (June 2018-Sept 2020)[19]

	Birr		**USD[20]**	
Household size	*At end-line*	*At enrolment*	*At end-line*	*At enrolment*

[19] Data from HelloCash Analytics.
[20] Exchange rate based on an average of monthly exchange rates of the period June 2018 to September 2020. Source: https://www.investing.com/currencies/usd-etb-historical-data, accessed 22/3/2021.

up to 2	516.49	422.11	16.92	13.83
3 to 5	489.05	493.13	16.02	16.16
6 and more	565.46	565.46	18.52	18.52

68. Irregularities in the monthly disbursement of mobile vouchers were captured in qualitative interviews both with beneficiaries and the implementation side. Reasons for this included network issues, errors during registration, and delays after changing registered phone numbers. IDI with beneficiaries further revealed a perceived inadequacy of the voucher amount based on family sizes and frequent price fluctuations of local fresh foods.

Social and Behaviour Change Communication (SBCC) Exposure

69. Next, we looked at the outreach of SBCC by awareness of different activities and participation at interactive events. We categorized awareness of SBCC activities by main events (cooking demonstrations, coffee conversations, radio shows, information material such as posters) and accompanying events (TV, loudspeaker, mobile SMS and calls, newspapers/magazines). As indicated, coffee conversations, cooking demonstrations, and HEW visits were suspended in early 2020 due to the COVID-19 outbreak. Nevertheless, more than half of respondents indicated that cooking demonstrations and coffee conversations took place in their community in the previous year. Together with radio shows and information material, 73% of respondents were aware of SBCC activities (Table 4). The low awareness of mobile SMS and calls, which have been rolled out as replacements for the interactive activities, may be explained by the fact that these have just started in the recent months.

Table 4: Households' awareness of SBCC activities

Awareness of	Number of households	% of households
Main SBCC activity	**974**	**73.0**
Cooking demonstrations	778	58.6
Coffee conversations	738	56.3
Radio shows	207	15.0
Posters/banners/boards/leaflets	182	13.2
TV	86	6.2
Local loudspeakers	14	1.0
Mobile SMS	11	0.8
Newspapers/magazines	10	0.7
Mobile calls	8	0.6
Any SBCC activity	**1001**	**74.9**

70. While interpersonal activities were suspended in early 2020, nevertheless 92% of those indicating awareness of coffee conversations or cooking demonstrations (or 55% of the total sample) have also participated in one in the past 12 months. Around 70% of respondents received a visit of a HEW, Health Development Army (HDA) or Women's Development Army (WDA) worker in the past 12 months. 81% of respondents recall 3 or more messages shared

about IYCF practices and PLW nutrition from HEW visits, cooking demonstrations, coffee conversations, and other SBCC activities[21].

Key findings and conclusions – Question 2.1

- With 97% of enrolled households meeting the eligibility criteria, the programme seems to target well the intended beneficiaries. However, only 20% are benefitting from PSNP support and almost three quarters of households no longer enrolled still exhibit eligibility, indicating that households graduate out of the programme before they become ineligible.
- While around 95% of enrolled respondents received a voucher in the past 6 months, only around 58% received vouchers every month. Results on average voucher amounts per household size are inconclusive but indicate that the disbursement was not regularly updated by the number of household members.
- SBCC activities reached 3/4 of respondents. While interpersonal activities were suspended in early 2020 due to the COVID-19 pandemic, still 55% of respondents participated in coffee conversations or cooking demonstrations in 2020. A small number (less than 1%) received messages through newly developed SBCC channels (mobile SMS and calls).

2.5. Evaluation Question 2.2: What are the effects of the project on availability of fresh foods?

71. Availability of fresh foods on the markets covered by the project was promoted by the programme through retailer engagement and training. Local traders offering a variety of fresh foods were identified as programme participants and trained in handling, procuring, and marketing high-quality fresh foods. To identify the supply and potential bottlenecks to the supply of various nutritious food, we first investigated the type and quantity of fresh food offered by traders.

Table 5: Number and % of traders supplying fresh foods

Traders offer at least	Number of traders	% of traders	Target 2020 (%)
Dark leafy greens	50	35.7	
Orange vegetables	47	33.6	
Orange fruits	64	45.7	
Animal-sourced food	16	11.4	
One of the categories	106	75.7	>80

72. Around 76% of the traders surveyed were able to offer at least one product from the food groups dark leafy greens, dark yellow/orange vegetables, dark yellow/orange fruits or animal-sourced food (Table 5). This is slightly below the target of 80% indicated in the logframe. However, it should be noted that the measurement of the indicator was updated for the end-line. Around 33% to 35% of traders offered orange vegetables and leafy greens, respectively, with a slightly higher amount of 45% offering orange or dark yellow fruits. 11% of the surveyed

[21] Information shared at cooking demonstrations and coffee conversations included diversification of porridge and meals for PLW. Potential messages received from HEW visits and other information channels included additional information on breastfeeding, introduction of complementary foods, meal frequency, and food preparation.

traders offered animal-sourced foods, among which the majority traded eggs, followed by meat and milk or dairy products.

73. Table 6 further shows that, while all traders indicate that they know where to procure their fresh foods all year around, 87% of traders are able to supply their fruits and vegetables throughout the year. IDI with traders suggest that seasonality plays a role for certain products (e.g. mangos or oranges), however this results in decreasing supply rather than unavailability. From Table 6, we further derive that the majority of traders have an understanding of how to handle fresh foods, based on training material related to waste measures, storage, and maintaining quality. The results exceed the target set in the logframe (however, again it should be noted that this indicator was updated at end-line).

Table 6: Traders who understand how to procure and handle fresh food

	Traders who	Number of traders	% of traders	Target 2020 (%)
Procurement and Supply	Know where to procure FF all year	137	100.0	≥75
	Supply FF all year	122	87.1	
Handling	Knows measures to reduce waste	133	95.0	≥75
	Knows measures to ensure quality	135	96.4	
	Store FF inside	124	96.9	

74. While 99.7% of beneficiaries indicated during the survey that they are either very satisfied (85.3%) or somewhat satisfied (13.8%) with the products bought with the FFV at the market, qualitative interviews revealed that quality of the offered fresh foods was an issue in several instances. The discrepancy could be because of one of two reasons: first, the quantitative findings only cover satisfaction with products that were bought, while qualitative interviews investigated the overall perception of offered products; and second, that the qualitative interview setting allowed for a more open and intimate discussion about experiences at the market. It emerged from the discussion with beneficiaries, traders, and market-level implementers alike that close follow-up by market supervisors with traders offering spoiled or low-quality foods was an important instrument that led to the elimination of the problem in many cases.

Table 7: Price levels and collusion

Traders who	Number of traders	% of traders
Offer at least 1 product with stable price level	51	36.4
Set prices: in coordination with other traders	22	16.1
Set prices: based on market or government prices	94	68.6

75. Price fluctuations played an important role in the access to fresh foods for FFV beneficiaries. Only around one third of traders offered products at a stable price in the past 12 months (Table 7). Qualitative insights indicate that unstable prices may have been caused by overall food price inflation and seasonality, but artificial price increases and collusion seem to drive prices specifically at participating FFV traders. While 68.6% of surveyed traders claim that they set their prices based on market and government price levels and 16.1% indicate coordination

with other traders as a driver for their prices (see Table 7), beneficiaries in IDI reported unreasonably high prices allegedly caused by tampering with weight scales and price fixing among FFV traders. Woreda officials and market supervisors confirmed that similar cases have emerged and were handled by exclusion from the programme.

Key findings and conclusions – Question 2.2

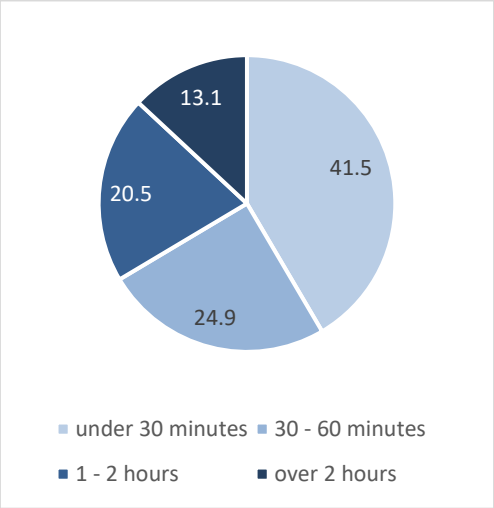
- Around 76% of FFV traders offer foods such as dark leafy greens, orange/dark yellow fruits and vegetables, or animal sourced products (eggs, meat, dairy products).
- While fresh foods can be procured and supplied throughout the year, seasonality plays a role in the amount and prices of fresh foods.
- Price instability and increases of FFV products seem to play an important role in the adequacy of voucher amounts. Strong supervision of market prices and retailer behaviour has been key to ensure that fair prices are established.

2.6. Evaluation Question 2.3: What are the effects of the project on access to fresh foods?

76. Access to fresh foods refers to the ability to use the digital vouchers and the possibility to redeem them at a market close to the beneficiary, in very rural areas. The first question was assessed through quantitative and qualitative findings. Out of the 1,114 beneficiaries indicating that they have received a voucher in the past 6 months, 96.1% confirmed that they have received training on how to access and redeem the FFV. This was further highlighted in IDI, where beneficiaries displayed confidence in their ability to use the voucher system on their own. Besides initial training at registration, additional home visits before the pandemic and assistance at the marketplace were applied to ensure the correct usage of the vouchers.
77. In light of the COVID-19 pandemic and subsequent lockdown measures following since early 2020, the main question of interest became to what extent beneficiaries had access to fresh foods throughout the pandemic. Beneficiaries were asked about the time necessary to reach the nearest market with their main means of transportation (including walking)[22]. The majority of respondents (42.5%) have a market within 30 minutes of reach where they can redeem their vouchers (Figure 1). Another quarter of respondents indicated a distance of 30 to 60 minutes, with 20.5% and 13.1% facing a distance of one to two hours or more than two hours, respectively.

[22] 54.4% of respondents identified walking as their main means of transportation, 22.8% respondent car/truck, 21.1% three-wheel bajaj, 1% animal-drawn cart, and below 1% motorcycle or other means such as horses or mules.

Figure 1: Distance to the nearest market (% of beneficiaries)



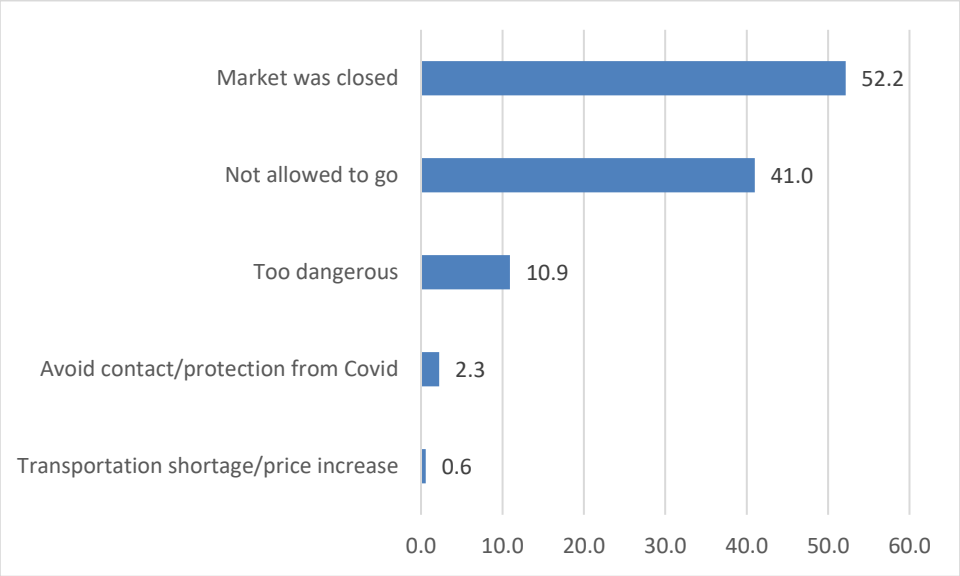
78. When examining the average distance to the nearest market by mode of transportation (Table 8), it becomes evident that the travel distance is closely correlated to the mode of transportation. While with a small vehicle, such as a motorcycle, animal-drawn cart or three-wheel bajaj, the average length of travel is 20 to 30 minutes, the average travel time increases for beneficiaries with a car or truck. Not surprisingly, the average distance is longest for those walking to the market with approximately 94 minutes on average. An increasing length of travel implies greater restrictions to market access and potential safety risks. Indeed, qualitative interviews have also revealed that long walking distances and lack of transportation are perceived as a challenge to both the beneficiaries, as well as the implementers at market level.

Table 8: Average distance to market by transportation mode

Transportation mode	**Average distance to market (minutes)**
Car or truck	42
Motorcycle	30
Animal drawn cart	20
Walking	94
Three-wheel bajaj	30
Total	**68**

79. The COVID-19 pandemic likely aggravated physical access to food markets. In fact, 34.5% of beneficiaries say they were restricted in their market access during the COVID-19 pandemic. Over half of those respondents indicating restrictions explained this with market closure, 41% indicated they were not allowed to go to the market (Figure 2). Around 10% mentioned it is too dangerous (without specification), while around 2% specifically indicated they did not enter the market to avoid contact and/or to protect themselves from a COVID-19 contraction. A small number of respondents named transportation shortages, e.g. due to sharp increases of transport costs, as a reason for limited access.

Figure 2: Reasons for limited access to market



Key findings and conclusions – Question 2.3

- Beneficiaries were informed and intensively trained on how to use the mobile FFV. Further assistance was available at the market.
- General access to markets by distance depends on the means of transportation and is mixed across beneficiaries. Results indicate that a majority of beneficiaries reaches a market where FFV could be redeemed within 30 minutes. However, average walking distances of over 1 ½ hours imply restrictions to market access.
- Markets were partly affected by the COVID-19 pandemic and subsequent lockdowns. Around one third of beneficiaries indicated that markets could not be accessed as usual during the pandemic.

2.7. Evaluation Question 2.4: What are the effects of the project on the demand for fresh foods?

80. The demand for fresh food is investigated by the frequency and amount of voucher redemptions and the types of products purchased. Further, we explored beneficiaries' knowledge and application of favourable infant and young child feeding practices to assess whether awareness-raising activities were successful in creating increased demand for fresh foods.

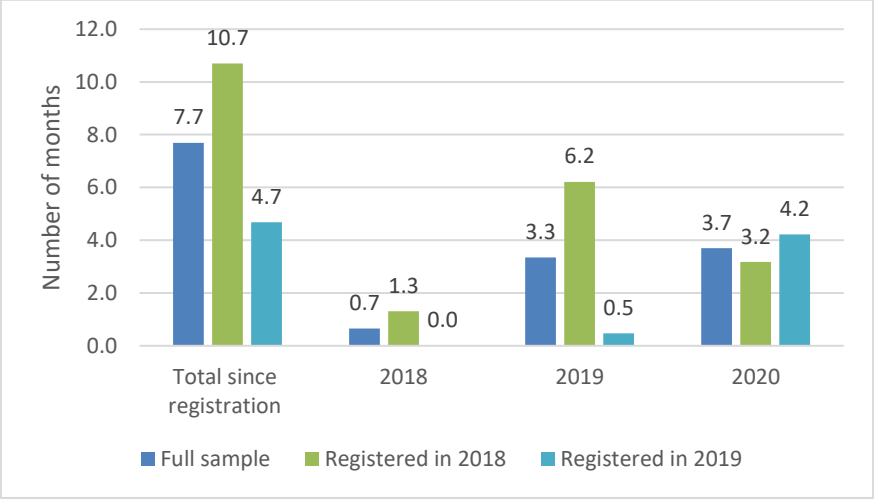
Voucher usage

81. Voucher redemptions were analysed using HelloCash data from June 2018 to September 2020 and cross-checked with end line survey data where possible. Of the sampled beneficiaries, 92.5% performed mobile money voucher transactions at least once since their registration. Compared to this, 99.8% of surveyed beneficiaries indicated they made use of the last voucher they received.
82. To examine the characteristics in which households with very few or no redemptions differ from households with frequent redemptions, we analyse statistical differences between the two groups. Table A 14 in Annex 8 shows statistics for households with 3 or fewer redemptions

(as a proxy for low voucher usage) and households with at least 4 redemptions. We find that households that hardly used the vouchers have smaller children (aged 6-23 months as opposed to 24-48 months) and generally less children living in the household, are less likely to have formal education, are fewer away from markets and are less likely to have electricity.

83. The frequency of redemption in terms of months is shown in Figure 3. When examining the total number of months in which redemptions were made since registration, the average amount of redemption months is 7.7. Due to the fact that those registered in 2018 had a longer period of time in which they could redeem their vouchers, the amount is largely driven by this part of the sample. To investigate this average further, we split up the total by redemption frequency in from2018 to 2020. In 2018 (where only those that were registered in 2018 could be redeeming vouchers), redemption frequency is low with 1.3 months on average. This may be explained with the registration and voucher disbursement having been slowly rolled out since June 2018. In 2019, the number of months in which those registered in 2018 have redeemed their vouchers can serve as an annual proxy: beneficiaries used their voucher in around 6 months, or on average every two months in 2019. Those in our sample enrolled in 2019 were only registered starting in August, such that for those households the number of months in 2020 is most telling in terms of frequency. Note that the data for 2020 only reaches until September, such that again a pattern of voucher usage every two months (4.2 out of 9 months) emerges. The lower number of months in which those registered in 2018 have redeemed their vouchers may be explained by a higher graduation rate in 2020 among early registrants.

Figure 3: Frequency of FFV redemption in months[23]



84. On average, beneficiaries redeemed their voucher 37 times in total. Spread across an average of 7.7 months, this implies an average of 4.8 redemptions per month. We cross-checked this information during the survey, where beneficiaries indicated that they redeemed their last voucher on average 3.5 times in the last month. The slightly lower amount may result from the shorter recall period or the simple fact that beneficiaries were not necessarily surveyed at the end of the month, or from more substantial issues such as continued restricted access to the market due to COVID-19 or lack of demand for the product.

[23] Data from HelloCash Analytics.

Table 9: Voucher value redeemed as share of total voucher value received[23]

	Share of value redeemed
Total	86.0%
By registration:	
Registered in 2018	81.1%
Registered in 2019	90.8%
By woreda:	
Dawa-Cheffa	93.6%
Dessie-Zuria	89.8%
Habru	77.8%
Kalu	89.9%
Kobo	75.6%
Mekdella	89.1%

85. The average voucher value redeemed as a share of the total voucher value received is on average 86% (Table 9). While this is an indication of high demand of the subsidised goods, earlier discussed statements about inadequate voucher amounts due to high prices may lead to expectations of even higher redemption shares. The data suggests that there are substantial differences in the used voucher volumes by (i) registration date and (ii) woreda. In particular, shares are higher for those registered in 2019 and those living in woredas such as Dawa-Cheffa, Dessie-Zuria and Kalu. With steady demand, this may imply increasing prices over time and across woredas. Contrarily, it could also be an indication of differential demand when steady prices are expected, i.e. due to increased awareness-raising or local demand patterns. To investigate these pathways further, more disaggregated data would be required.
86. When looking at the types of products bought by beneficiaries, 96.5% of vouchers were redeemed at fruit or vegetable shops, 2.5% at egg traders, and the remaining share at meat (0.9%) and milk shops (0.1%). By product, 66% of traders offering dark leafy greens indicated they sold most or all of their produce to FFV beneficiaries; similarly, 71.7% of yellow/orange vegetables and 69.8% of yellow/orange fruits were mostly or entirely sold to FFV households (see Table 10). For animal sourced products, this number reduces to 26.7%.

Table 10: Percent of traders selling most or all of their produce to FFV beneficiaries

Type of Product	Number of traders	% of traders[24]
Roots and tubers	85	72.6
Nuts and pulses	1	100
Dark leafy greens	33	66
Yellow/orange vegetables	33	71.7
Other vegetables	78	72.9
Yellow/orange fruits	44	69.8
Other fruits	73	84.9
Animal sourced food	4	26.7

87. Looking at the consumption patterns of women and children (Table 11), we see that the most consumed food groups by women are grains, roots and tubers, pulses and seeds, and other

[24] Percentage based on traders selling the respective product.

fruits and vegetables. Dark green leafy vegetables and other vitamin A rich foods such as orange vegetables and fruits are consumed by around 36% and 20%, respectively. Besides grains and roots, the most consumed food groups by children are vitamin A rich foods (53.3%), eggs (52.2%) and other fruits and vegetables (47.2%).

Table 11: Consumption of women and children by food group

Food Group	% of beneficiaries	% of children
Grains, roots and tubers	99.6	91.0
Pulses and seeds	85.8	37.4
Dark green leafy vegetables	35.7	29.1
Other vitamin A rich foods	20.4	53.3
Other fruits and vegetables	71.3	47.2
Dairy products	16.4	43.3
Meat	5.3	3.9
Eggs	14.5	52.2

Knowledge and IYCF Practices

88. As a driver for demand for fresh foods, knowledge about IYCF practices was investigated through the survey. Figure 4 shows that 94.1% were aware of the correct age to start complementary feeding (6 months). Around 98% further knew reasons to start complementary feeding (e.g. to help the child grow or gain strength) and what kind of foods to complement to breastfeeding (e.g. porridge, fruits and vegetables). Vitamin A rich foods were less known by 53.3% of respondents. The number of correct reasons to start complementary feeding was on average 2.6, while around 4 complementary foods and 1.5 vitamin A rich foods were identified (Table 12).

Figure 4: Percent of beneficiaries with knowledge of IYCF practices

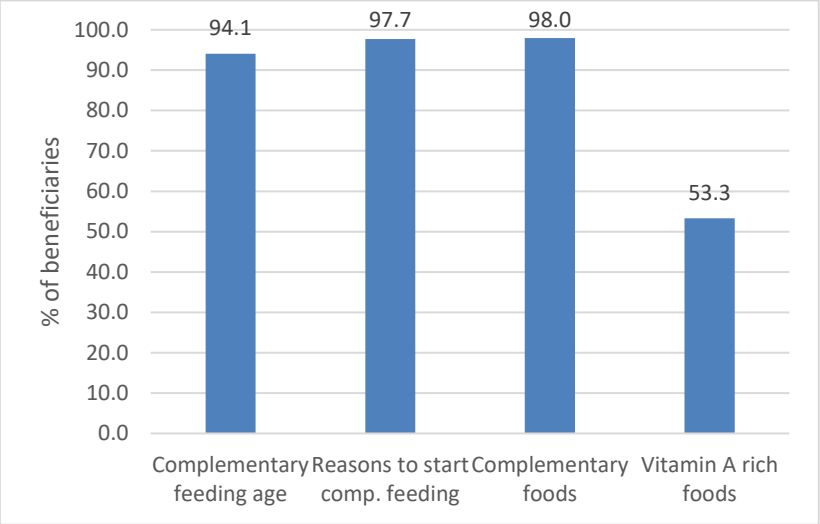


Table 12: Average number of known answers per question

Knowledge Question	Average correct answers
Reasons to start comp. feeding	2.6
Complementary foods	4.2
Vitamin A rich foods	1.5

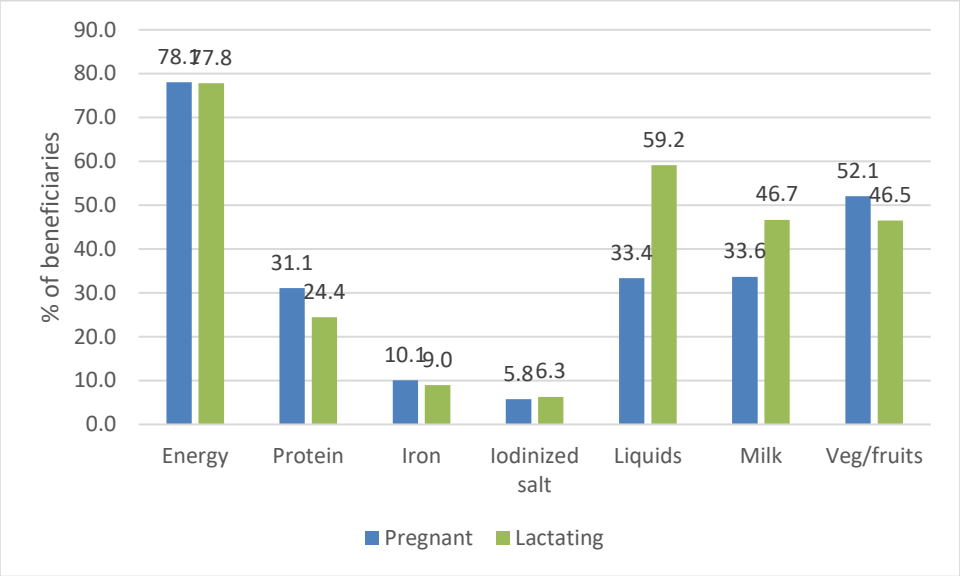
89. Knowledge about healthy nutritious practices for PLW are also widely prevalent: 95.5% and 96.8% know at least one important change to the diet when pregnant or lactating, respectively. Both topics are known to 94.4% of the respondents (Table 13).

Table 13: Knowledge of changes in nutrition for PLW

Knowledge of:	% of beneficiaries
Nutrition for pregnant women	95.5
Nutrition for lactating women	96.8
Nutrition pregnant and lactating women	94.4

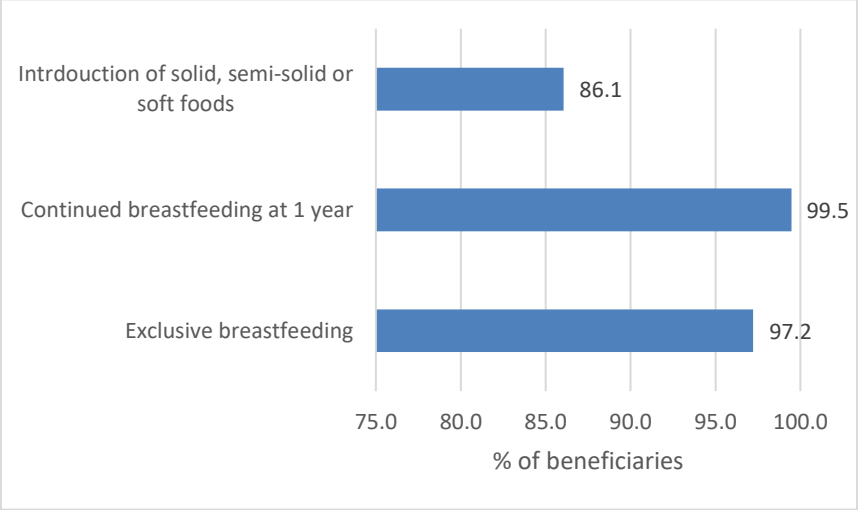
90. As Figure 5 shows, respondents were most aware about the need for higher energy intake during pregnancy and lactation. Other well-known topics were increased servings of fruits and vegetables (52% for pregnant and 46% for lactating), increased liquid intake and serving milk. The latter two, however, were wider known in case of lactation. Contrarily, the importance of eating more protein and iron-rich foods was slightly more understood for pregnancy. Overall, few respondents were aware of using iodized salt when preparing meals for pregnant or lactating women.

Figure 5: Type of PLW diet change known



91. When examining which IYCF practices were applied by respondents, we found that 97.2% out of 72 respondents with infants aged up to 5 months followed exclusive breastfeeding, 99.5% out of 193 respondents with children aged 12-15 months continued to breastfeed their children, and 86.1% out of 192 respondents with infants aged 6-8 months introduced solid, semi-solid or soft foods to their child's diet (see Figure 6).

Figure 6: IYCF Practices applied by beneficiaries



Key findings and conclusions – Question 2.4

- Demand for fresh food seems to be high: beneficiaries make regular use of their vouchers throughout the year and several times per month. On average, 86% of the voucher amount is redeemed.
- Beneficiaries mainly redeem their vouchers at fruit and vegetable traders. A majority of traders sell a large portion of their products to FFV beneficiaries. Main food groups

consumed by PLW and children include grains, roots and tubers, pulses and seeds, vitamin A rich foods and other fruits and vegetables.

- Knowledge of healthy IYCF and PLW diet practices is generally high among beneficiaries. However, knowledge of vitamin A rich foods and benefits of protein and iron-rich foods are lesser known.
- IYCF practices such as exclusive breastfeeding until 6 months and continued breastfeeding at 1 year are widely applied, however introduction of complementary foods at 6 to 8 months is less common among respondents.

2.8. Evaluation Question 2.5: Did the intervention produce the expected nutritional results?

92. In this subsection, we present statistics on the main nutrition outcomes of the programme, namely the Minimum Acceptable Diet (MAD), the Minimum Diet Diversity (MDD) and the Minimum Food Frequency (MFF) for children, as well as Minimum Diet Diversity outcomes for mothers (MDD-W). These are compared to baseline (2018) and midline (August and December 2019) values. In addition, we present multivariate regression estimates for each outcome to examine the mechanisms and parameters, by which each nutrition outcome is affected. We further present comparisons between children aged 6-23 months in active households with children aged 24-48 in active households, to examine possible spill-over effects for older children in enrolled households.
93. We restrict the analysis of nutrition outcomes to only those living in households that were still enrolled in December 2020, i.e. at the time of end-line data collection. This implies that the analysed sample reduces to 794 (54.5% of the total sample).
94. As mentioned before, no anthropometric measurement was possible because of COVID-19. Hence, only reported food consumption can be analysed. Further, as the programme focuses on enhancement of dietary diversity rather than direct nutritional results, these reported measures seem to be the most appropriate for the analysis.

Nutrition scores for children

95. Table 14 shows end-line nutrition scores for children aged 6-23 months in households that registered in 2018 and 2019 and have been enrolled as of December 2020. The table further reports past nutrition outcomes from the baseline (pilot) in 2018, from Post-Distribution Monitoring (PDM) data collection in August and December 2019, as well as the target outcomes for 2020 as reported in the logframe. The presented outcomes are the MAD, MMF and MDD[25]. We find that an average of 42.9% of children in our sample meet the MAD criteria, hence reaching the 2020 target. When breaking down the indicator into MDD and MMF, we find that 46.5% of children meet the MDD criteria.

Table 14: Nutrition outcomes for children aged 6-23 months (enrolled in December 2020)

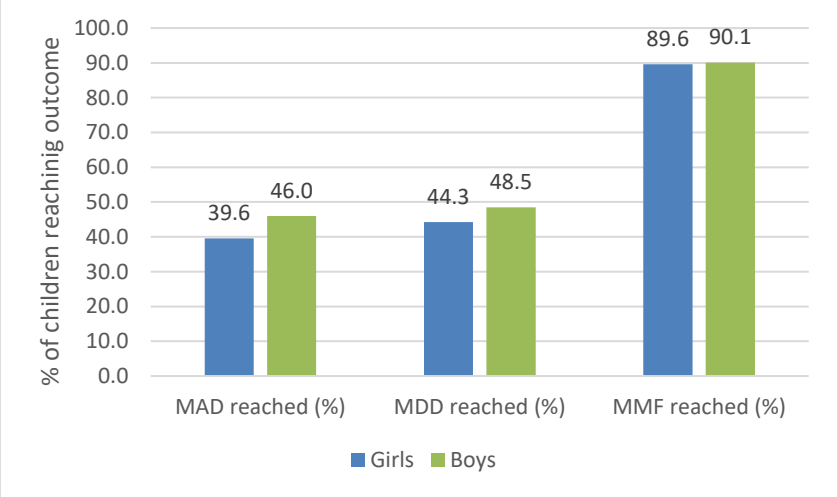
Outcome indicator	2018	2019	2020

[25] To calculate the MDD, we first define 7 food groups according to WHO guidelines, including grains, roots and tubers (1), pulses and nuts (2), dairy products (3), flesh foods (4), eggs (5), vitamin-A rich fruits and vegetables (6) and other fruits and vegetables (7). We then define a dichotomous variable for the MDD, which is =1 if the child has eaten at least four of the food groups, and 0 otherwise. To measure the MMF, we define a second dichotomous variable which is =1 if a child of 6-8 months is breastfed and gets food at least two times per day (1), if a child of 9-23 months is breastfed and gets food at least three times per day (2) and if a child of 6-23 months that is not breastfed gets food at least four times per day (3). The MAD is defined as a dichotomous variable that is =1 if a child of 6-23 months meets the requirements of both MDD and MMF.

	Baseline	August	December	Target	**End-line**	
	%	%	%	%	%	N. Obs.
% Meeting MAD	22.3	40.2	93.3	42.3	**42.9**	692
% Meeting MDD	22.3	59.6	96.6	42	**46.5**	706
% Meeting MMF	86.8	63.2	96.6	90	**89.9**	631

96. There is an increase across all scores compared to 2018 levels and the targets for 2020 have been reached. However, compared to December 2019 levels in 2020 are substantially lower. Several reasons may play into this discrepancy: first, while the indicator composition was harmonized, the sampling and data collection methods between PDM and end-line differ substantially. Further, we cannot measure the levels that would have been reached without the COVID-19 pandemic.

Figure 7: Nutrition scores by gender



97. Figure 7 shows nutrition scores of children by gender. In our sample of enrolled households, we find that girls have slightly lower scores for all three outcomes. Still, the MDD target is reached across gender, while for MMF and MAD the target is reached only for boys.
98. While the main target of the programme are children aged 6-23 months, we further investigate the hypothesis that older children in enrolled households may benefit from potential spill-over effects, i.e. that older siblings above the age of 2 still benefit from the program.

Table 15: MDD by age group

Age group	**% of children with MDD=1**	**Total number of children**
6-11 months	36.6	257
12-23 months	52.1	449
24-35 months	38.3	81
36-48 months	47.1	153

99. Table 15 shows percentages of children who met MDD requirements for age groups 6-11 months, 12-23 months, 24-35 months and 36-48 months in enrolled households. We find that children of 12-23 months have the highest likelihood to meet the MDD score. However, for older siblings the share of children reaching the MDD drops for two- to three-year-olds and increases again with age. We thus find inconclusive results in terms of spill-overs within households that cannot be further analysed given the lack of a control group.

100. When breaking down the food groups constituting the MDD (Table 16), we find that the four main food groups consumed by children are grains, roots and tubers, followed by vitamin-A rich foods, eggs, and other fruits and vegetables. When comparing the diet of boys and girls, it becomes evident that while the main food groups are same across gender, girls are somewhat less likely to receive eggs and other vegetables and fruits, but more likely to receive vitamin A rich foods.

Table 16: MDD Food Groups consumed (by gender)

Food Group	**% children (total)**	**% children (boys)**	**% children (girls)**
Grain, roots and tubers	94.5	94.2	94.7
Pulses and nuts	27.2	28.1	26.2
Dairy Products	40.3	40.4	40.1
Flesh foods	3.5	3.9	3.0
Eggs	57.9	60.3	55.2
Vitamin A rich foods	61.5	60.3	62.9
Other fruits and vegetables	53.1	54.7	51.3

101. To further examine mechanisms that influence each respective nutrition outcome, we run several multivariate regressions, including a number of possible explanatory factors at the level of children, household and market. The regression model, as well as results are presented in Annex 9. The regressions indicate that the MAD is significantly determined by age of the child, but with diminishing effects as age increases (i.e., one additional month of age has a less strong effect on older children than on younger children). Considering age coefficients for MDD and MFF, it appears that the impact of age on the MAD is mainly driven by the MDD. These insights may indicate behavioural aspects of child feeding, i.e., that older children receive a wider variety of foods. Further follow-up would be needed to understand the underlying knowledge or attitudes towards child feeding at different ages.

102. Among household characteristics, the mother being married is significantly associated with higher likelihood of the child meeting MAD requirements. While being married has a negative effect on the MMF, it has a positive effect on the MDD. Further, children that were breastfed at the time of interview are more likely to meet MAD requirements, but this effect is not statistically significant for the MDD and MFF outcomes.

103. Another positive determinant of MAD, MDD and MMF scores is the household owning milk cows, while more children in the household decrease the likelihood for a given child to meet MAD requirements. We do not find any significant associations of PSNP transfers or reception of fresh food vouchers with the MAD indicator.

Nutrition scores for women

104. Table 17 shows end-line scores for the MDD-W for mothers aged 18-49 years in enrolled households that registered in 2018 and 2019[26]. Analogously to Table 14 in the child nutrition section, this table further reports past nutrition outcomes from the baseline data collection in 2018, from PDM data collection in August and December 2019, as well as the target outcomes for 2020.

Table 17: Nutrition outcomes for women aged 18-45 (enrolled in December 2020)

Outcome indicator	**2018**	**2019**		**2020**		
	Baseline	August	December	Target	**End-line**	
	%	%	%	%	%	N. Obs.
% Meeting MDD-W	3	70	93	50	**31.7**	712

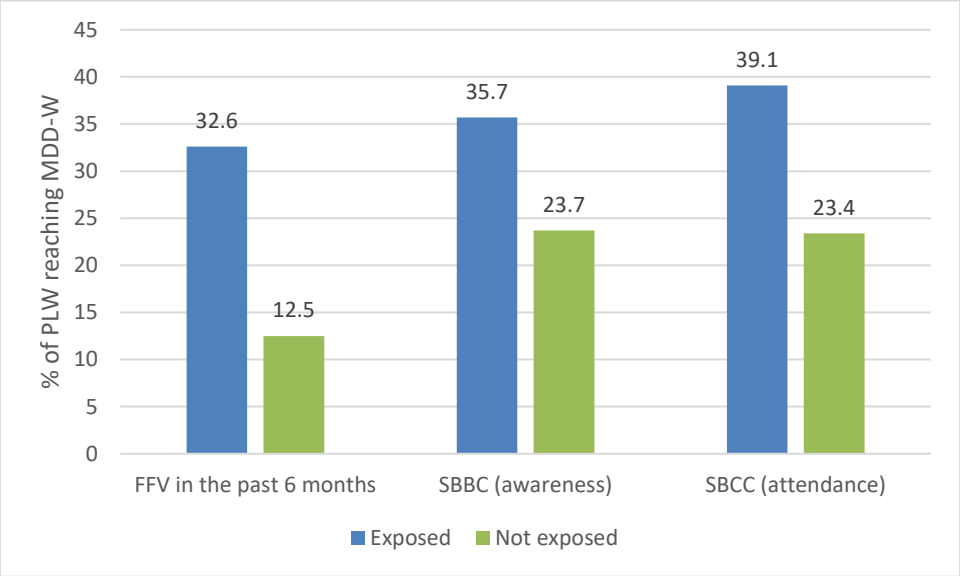
105. We find that among women in our sample, 31.7% meet the MDD-W criteria, missing the 2020 target of 50% by 18.2 percentage points. Furthermore, the outcome is lower than the PDM outcome of 70% in August 2019 and the outcome of 93% in December 2019. However, compared to the low value at baseline we see a substantial increase in the share of women reaching the MDD-W score.

106. Again, given the different sampling methods, as well as the unforeseen circumstances surrounding the pandemic, it is difficult to reach meaningful conclusions by comparing the results. One factor that was investigated was the potential influence of the Orthodox Christian fasting period at the time of the survey (i.e. no animal products are consumed, and the first meal is delayed until the afternoon). However, while we did not collect information about religious affinity, we believe religious fasting to be an unlikely driver of low dietary diversity based on the survey regions, which are predominantly Muslim areas. Further review of "normal" eating habits in the previous day did not reveal any additional insight. In fact, respondents who indicated an event having affected their normal eating habits reported more consumption of meat, while consumption of eggs and dairy products was lower for this group.

107. Given the low outcomes, we aim to disentangle the drivers of the lower score by looking more closely into the exposure of PLW to the programme components.

[26] MDD-W is defined based on FAO guidelines as follows: The indicator is equal to 1 if 5 or more out of 10 food groups were consumed by the woman within the last 24 hours. The food groups are (1) grains, roots and tubers, (2) pulses, (3) nuts and seeds, (4) dairy products, (5) meat, poultry and fish, (6) eggs, (7) dark green leafy vegetables, (8) other vitamin-A rich fruits and vegetables, (9) other vegetables, and (10) other fruits.

Figure 8: Percent of PLW reaching MDD by programme exposure[27]



108. As indicated in Figure 8, we find that 32.6% of women who indicate that they received a voucher in the past 6 months meet the MDD-W requirements, as compared to 12.5% of women who did not receive any FFV vouchers in the past 6 months. We further find that 35.7% of women who have received messages through SBCC activities meet the MDD-W requirements. The highest share of women reaching the MDD-W score is found within those that have attended cooking demonstrations or coffee conversations. These results indicate that having received a voucher and having been involved in SBCC outreach play an important role in increasing the dietary diversity for women.

109. To further examine possible mechanisms and characteristics by which the MDD-W outcome for women may be affected, we implement a multivariate regression including a number of explanatory factors at the level of household and market, similar to the regression model for children's nutrition above. The full model and results are described in Annex 9. From the regression, we find that women whose youngest child is between 24 and 48 months old are significantly less likely to meet MDD-W requirements compared to women whose youngest child is less than 6 months old. Furthermore, the total number of children in a household significantly decreases a woman's likelihood to meet MDD-W requirements. Women in woreda Kobo are significantly less likely to meet MDD-W requirements. Further, we find that women in households that own oxen or bulls are significantly more likely to meet MDD-W requirements. The likelihood of a woman meeting MDD-W requirements decreases significantly if the household faced restricted access to markets during the COVID-19 pandemic. Having received an FFV in the last 6 months significantly increases the likelihood of women meeting MDD-W requirements.

Limitations to remote data collection

110. It should be pointed out that collecting nutritional data remotely via phone interviews is challenging in several ways. Firstly, as pointed out before, anthropometric measures of mothers and children could not be gathered. These measures would allow for an objective

[27] Statistics based on enrolled households.

assessment of nutritional status. On the contrary, interviewing about food consumption may be prone to bias due to recall errors or social desirability.

111. These issues may have been exacerbated by the remote data collection. In particular, due to the time constraints of the phone interviews, food consumption was not asked in a two-stage process as is common practice, i.e., the respondent is first asked to list all meals they had during the day or night, which are then categorized into the food groups. Rather, a combined question was asked about the food groups consumed for each meal.

112. While these limitations persist, they were mitigated by an intensive training of enumerators to understand and categorize food groups correctly based on the respective meals. Further, answer options included exemplary meals and foods that contain the respective food groups. Enumerators probed for any food or drinks consumed throughout the last 24 hours, as well as for all ingredients of the dishes consumed.

Key findings and conclusions – Question 2.5

- Compared to baseline values in 2018, nutrition outcomes for children increased at end-line; however, they decreased compared to the December 2019 values. Targets for 2020 are nevertheless met for MAD, MDD and MMF.
- The MDD target is reached across gender, while for MMF and MAD the target is reached only for boys.
- Age plays a role in the MAD and MDD score, as they increase with an additional month of life for children. Other relevant drivers of the MAD are the marital status of the mother, children being breastfed at the time of the interview, and the household owning milk cows. Children with less siblings have a higher MAD score.
- Targets for minimum MDD-W in 2020 were not met and are lower than December 2019 values, however they have substantially increased from the baseline.
- Having received a FFV and having been exposed to SBCC seem to play an important role for mothers' MDD-W score.
- Further drivers of MDD-W include the age of the youngest child (with a higher age being increasing MDD-W), the number of children in the household, livestock ownership, and the access to a market during the COVID-19 pandemic. Further, women living in woreda Kobo are significantly less likely to meet MDD-W requirements, even after controlling for a set of household variables.

2.9. Evaluation Question 2.6: Is there an increase of financial inclusion and financial autonomy among female beneficiaries?

113. Financial inclusion and autonomy of women were secondary aims of the programme, that is, of the digital system of transfer. Potential channels could be that (i) women are given the power to make food purchases for the household through the voucher and, thereby, shifting intra-household decision making; and that (ii) women gain more financial independence by making use of the mobile money wallet for other services. As the programme did not foresee direct monitoring of who in the household made use of the mobile money wallet, we aimed to reach insights into this aspect mainly through qualitative tools.

114. The IDI with beneficiaries revealed a sense of female empowerment both in purchase decisions and the familiarization with the banking system. A majority of respondents claimed to be mainly or equally in control of the use and redemption of the vouchers. Through awareness creation and training in usage of the mobile money wallet, the beneficiaries gained

confidence in financial decision making. As explained by a woreda coordinator, beneficiaries were encouraged to self-help by building groups:

"In the beginning, some beneficiaries were not able to check their balance before and after they make the shopping. Later, we grouped them into (groups of) 1 to 10, and then those who can write and read help those who are illiterate. A beneficiary who reads and writes orient/show them how to check balance, transfer money and the like." (Woreda coordinator, male, Sekota)

115. Nevertheless, less than 10% of surveyed women stated that they make use of the mobile money wallet for other purposes than the FFV, indicating that other mobile banking services are not being used more as a consequence of the programme. This is consistent with the fact that the voucher system set up by the programme does not allow to procure anything else than fresh produce. Another issue, particularly at the beginning of the programme, seemed to have been that beneficiaries were often registered with someone else's (e.g. their husband's) phone, potentially restricting the freedom of action of the female beneficiary. However, the qualitative findings suggest that most beneficiaries later re-registered an alternative phone which they had free access to.

116. As a majority of FFV traders were female, the programme may have further implications for financial independence of female traders. Besides offering economic opportunities, the traders are trained in handling the HelloCash mobile system:

"At the time we were afraid of it. When they give us the training, they showed us how it works. Then they come and give us training on how it works, how the money enters, how we get the money once it is transferred. (Trader, female, Dessie Zuria)

Key findings and conclusions – Question 2.6

- The programme did not specifically monitor access of the voucher by female beneficiaries. Consequently, financial inclusion and autonomy are potential side benefits resulting from the programme.
- Qualitative evidence suggests that women (both beneficiaries and traders) gained a sense of empowerment and control due to awareness raising and training in usage. Purchase decisions seem to have been primarily made by women.
- Limited evidence suggests that the programme had no effect on the use of financial services by women.

2.10. Evaluation Question 2.7: Was the intervention efficient compared to possible alternatives?

117. The FFV programme combines several innovations compared to traditional alternatives, aimed at increasing diet diversity in mothers and children. The approach is expected to save distribution costs as compared to food provision in in-kind and offer digital vouchers. Engagement of private sector actors has the potential to make the inclusion of small retailers and fresh food supplier and takeover more effective.

118. Alternative approaches used by WFP in the region include direct food aid, cash-based transfers, and traditional paper-based food vouchers. The programme aimed to combine the advantages of these intervention types, namely the controllability of the type of commodities received and the stimulation of local markets, by offering a digital voucher redeemable for a

specific list of food items at local retailers. In terms of costs of the programme, the savings of food transportation and allocation are redirected into establishing digital voucher systems[28].

119. In Ethiopia, where cash transfer and food aid are well-known and common nutritional interventions, the use of electronic cash transfer using mobile technology in the most rural part of Ethiopia was shown to be feasible which has not been done any other nutritional interventions in the country. As a respondent from the government side stated:

"In Ethiopia, we like to pay cash on hand… but with the implementation of the WFP FFV programme the first year, it has become very clear that electronic cash transfer work for the vulnerable population is really feasible, even if very remote, areas… as long as the internet connectivity is there." (Seqota Declaration officer)

120. To ensure that only the intended products are bought, the programme entails constant monitoring of market activities and follow-up with suspicious voucher sizes redeemed. These tasks are assumed by local implementers in combination with customer support. For the next phase of the intervention, an innovation of the digital voucher system is envisioned, integrating an itemization of purchases by type, quantity and costs of bought products. If implemented successfully, this may improve effectiveness of the programme by standardizing the monitoring process of food purchases.

121. Compared to traditional voucher systems, a shortfall of the digital nature of vouchers is the dependence on phone ownership, network stability, technical capacity of the beneficiaries and novel monitoring requirements. Indeed, technical challenges in the registration and follow-up especially at the start of the programme were mentioned in KII. While these issues have been continuously addressed throughout the implementation phase, it remains unclear to what extent phone ownership restricts the programme from reaching those beneficiaries most in need of the intervention. This shortcoming should be weighed against the potential to reduce inefficiencies in the distribution and monitoring of vouchers.

122. Another innovative feature of the programme is the involvement of multiple unconventional stakeholders. Partnerships with the private sector are particularly crucial to the programme design, as it involves the banking sector and smallholder retailers alike. This engagement may play an important role in the co-creation and takeover of the intervention. Nevertheless, qualitative findings suggest that effectiveness of this approach may be hampered by the need for a strong supervision of the implementation at market level, as well as a level of mistrust of beneficiaries towards the private parties.

123. Lastly, the integration of SBCC with the disbursement of vouchers was viewed to add to the effectiveness of the programme as awareness of the utilization and nutritional rewards of fresh foods stimulated the demand for fruits, vegetables, and eggs.

Key findings and conclusions – Question 2.7

- The FFV programme combines several innovations compared to traditional alternatives, aimed at overcoming typical inefficiencies. The approach is expected to save distribution costs as compared to food aid and revolutionize the disbursement of vouchers.
- With this novel approach, lessons have been learned and incorporated along the implementation period and for future programming.

[28] No costs of the programme have been shared with the evaluation team. Statements are solely based on KII information and own elaborations.

- Engagement of private sector actors has the potential to make stakeholder inclusion and takeover more effective; however, dynamics between the private sector and beneficiaries should be taken into account.

2.11. Evaluation Question 3.1: What is the government readiness to take over the programme and what are the anticipated ownership structures?

124. Key informants indicated that there has been continuous government engagement on all levels, including federal, regional, zonal and market level. On the federal level, the primary counterpart has been the Seqota Declaration Senior Officer, represented throughout the programme design and implementation. Amhara Regional health officials and other regional offices were part of the supervision and follow-up of implementation.

125. Currently, the expansion of the programme is planned for 15 additional woredas with the ultimate goal to align the beneficiary targets to the Seqota Declaration, i.e., one million beneficiaries by 2025. The realization is planned by a coalition among WFP, the Ethiopian government represented by the Seqota Declaration, and all the other development donors involved in nutrition programming in Ethiopia. funding and support negotiations are ongoing with donors such as KfW, IFAD, USAID and others.

126. On the federal governmental level, through the coordination of the Seqota Declaration several agencies are involved in the future expansion including Ministries of Agriculture, Health, Finance, and Education.

127. Further, the national PSNP programmes have been informed by lessons learned from the implementation of FFV in the seven woredas of Amhara Region. Parts of the FFV programme are planned to be mainstreamed within the existing PSNP programme and the Seqota Declaration which includes 40 districts.

128. A high level of engagement has also been identified at the woreda level, including agriculture, health, and administration offices. The FFV programme functions under a nutritional task force that runs the programme at the district level led by NNP. Especially during the pandemic, the Trade and Industry Coordination Office played a great role to facilitate the allocations of the marketplaces for consistent purchasing and safety of interaction to prevent the transmission of the virus. The district health office participates starting from registration until the programme closes.

129. A task force from the district agriculture office, health extension office, women and children affairs office, and trade community office is responsible to follow, organize, facilitate and control the market activities every week. If problems cannot be solved by the members of the team, they report to the programme coordinator at the district level. If it is beyond the capacity of the district level office, it will then be taken to WFP through the health office in the district. The tasks include orientation and assistance for beneficiaries at the market and supervision of market activities.

130. Private-public partnerships include the Anbessa Bank as the service provider of electronic cash transfers and other banking services. Further, fresh food suppliers which are registered as small or medium scale business enterprises are private stakeholders to the programme. These enterprises are suppliers of fresh food to the traders in the woredas.

131. In terms of capacity building, the technical wing of Ethiopian Ministry of Health (EPHI) has been provided, by WFP, with comprehensive training on the “Fill the Nutrient Gap” analysis: an

approach that has the potential to strengthen nutrition situation by helping to identify the proper voucher value based on the market of fresh food. The FNG has informed the update of the FFV transfer value in 2019, leveraging its component called the Cost and Non-Affordability of a nutritious diet. MoH/EPHI was trained to adjust the FFV value based on the Cost / Non-affordability of nutritious diets, via the FNG method, now mainstreamed with technical support from WFP.

132. This programme was a pilot to generate evidence and inform future Ethiopian nutrition policies. Financial, as well as skill and technology needs were identified for the successful continuation of the FFV programme. As the programme has been heavily donor-dependent, support from diverse development stakeholders is needed until the government fully takes over the implementation. Discussions with the Ethiopian ministry of finance have been initiated through the Seqota declaration senior officer such that the government takes part in funding the programme.

Key findings and conclusions – Question 3.1

- The programme combines a high level of government engagement on federal, regional and woreda level throughout the implementation period.
- An expansion of the programme in additional woredas is planned and coordinated among government and other development stakeholders.
- Ownership structures further include private-public partnerships with the financial sector and fresh food suppliers/retailers in rural markets.
- Government capacities have been fostered in skills including nutrient gap cost and non-affordability of nutritious diet analysis to inform the re- adjustment of the FFV transfer value, and the contents of the SBCC implementation.

2.12. Evaluation Question 3.2: Is the programme capable of overcoming future challenges and bottlenecks?

133. Network issues pose a major challenge to the programme due to the digital nature of both voucher transfers and redemption at the market. Qualitative interviews with beneficiaries and traders revealed that transactions at the market can be affected if beneficiaries cannot transfer their payment right away due to network instability. This was solved on the market level by registering names and phone numbers of customers such that the beneficiary could receive the product on the condition to pay for them when network allows. Such practical solutions can be facilitated by the implementers at the market.

134. Further, network problems can cause delays in the disbursement of transfers to beneficiaries. With unreliable transfers, beneficiaries may be unaware of the receipt of a transfer and have less opportunities to redeem the voucher. Traders have also noted during IDI that it hampers their ability to supply adequate amounts of food when the time of disbursement is irregular. A mechanism to accommodate this situation was to sensitize beneficiaries to redeem their vouchers once network is back and inform traders of when transfers have been disbursed.

135. Besides delayed transfers due to network issues, another challenge emerged at the beginning of the programme in the timely disbursement of transfers after registration due to capacity constraints when large amounts of registrations had to be processed. This implies a need for adequate resources and capacities to handle the registration process given a certain

number of expected registrants. In line with this, mechanisms to ensure structured updating of the databases and re-registration of beneficiaries are required. IDI interviewees reported the need to change their phone number after registration as they had often initially used someone else's phone. This reportedly caused issues in receiving transfers following the re-registration.

136. Continuous follow-up and assistance were crucial for some beneficiaries as they initially did not have an adequate understanding of how the overall FFV system worked. Woreda coordinators assisted in informing about the available amount of money, which products were covered, and the maximum value allowed to purchase foods during one interaction.

137. Generally, market supervision and assistance played a crucial role to overcome challenges faced in the implementation. Continuous follow-up was needed not only for beneficiaries, but also to ensure adequate prices and quality of the fresh food products offered by traders. Traders who did not follow the rules, e.g. by artificially increasing prices above market level or selling low-quality products, were rebuked or, if continuous mishandling was observed, excluded from the programme. The strong need for constant supervision implies that traders should be increasingly incentivized to follow operating guidelines by themselves.

138. Insufficient voucher sizes, mainly due to the fact that a large percentage of the FFV beneficiary households (80%) were not receiving the PSNP transfer, high prices on the market and high demand for fresh foods, were predominantly considered a challenge for beneficiaries. To ensure an adequate voucher amount, a nutrient gap analysis cost / non-affordability of a nutritious diet was conducted at the beginning of the programme to inform the FFV transfer value adjustment occurred in 2019 and 2020. While this is a crucial approach, continuous price fluctuations imply a need for frequent assessments of local market prices and potential adjustments to the voucher size.

139. Other challenges mentioned include large walking distances to markets for both beneficiaries and market committees, which aggravates the possibilities to supervise and support market interactions. A need for enhanced transportation opportunities is identified.

Key findings and conclusions – Question 3.2

- Main challenges during the programme period included network issues, delayed transfers especially after registration, fostering understanding of the digital voucher system, and high needs for supervision to ensure fair prices and quality.
- There has been continuous improvement of mechanisms to address these bottlenecks. Lessons learnt have been incorporated in the registration process and market-level support.
- Remaining issues include the high need for market supervision, frequent price fluctuations affecting voucher adequacy, and lack of transportation to markets.

2.13. Evaluation Question 3.3: Will behavioural changes related to nutrition last after the programme?

140. Lastly, we aim to identify if nutrition behaviours change sustainably even after the household graduated from the programme. Clearly, within this evaluation we can only investigate nutrition behaviour in a very short term as the programme was ongoing at the time of the survey and households have graduated the programme only very recently. Longer-term

behavioural and nutritional changes can only be assessed several years after leaving the programme.

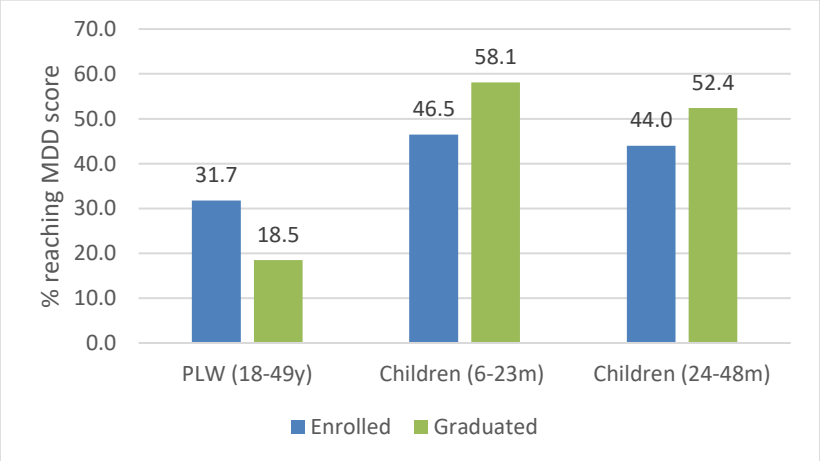
141. Qualitative evidence suggests that there is a shift in awareness about the benefits of fresh foods that drive continuous demand even when vouchers are no longer received. However, 80% of the FFV beneficiary households, although eligible, do not receive the PSNP core transfer cash, plus prohibitively high prices restrict households to purchase sufficient amounts of fresh foods without the voucher. Hence, mothers may face a trade-off between feeding nutritious fresh foods to children as opposed to consuming the products themselves.

142. To investigate this further, we looked at the nutritional status of women and children in graduated households as opposed to enrolled households. At the time of the end-line survey, 784 (or 54.5%) of households were enrolled in the programme, while 586 (or 42.5%) were no longer registered. Out of these, 40.5% were still active beneficiaries in October 2020, implying that they graduated from the programme within the last 3 months.

143. As noted above in EQ2.1 and shown in Table 1, almost three-quarters of graduated households seem to still be eligible for the programme. Table A 17 in Annex 9 demonstrates that 6.5% of children in graduated households are below 6 months and 31% are between 6 and 23 months old. This implies that the household has left the programme based on other reasons. For the following analysis, we therefore look at children between the ages of 6 months and 4 years.

144. As Figure 9 shows, 31.7% of women reach the MDD-W score in enrolled households, as compared to 18.5% in graduated households. This drop implies that women reduce their diet diversity once they no longer receive the programme. Conversely, we find that both younger (6 to 23 months) and older (24 to 48 months) children experience an increase in MDD when the household graduates from the programme. This relationship may indicate that women reduce their intake of nutritious fresh foods while keeping their children's consumption stable when no longer receiving the vouchers.

Figure 9: Percent of PLW and children in enrolled and graduated household reaching MDD



145. It should be noted that children in different age groups in enrolled and graduated households may differ in other ways and are therefore only imperfectly comparable. For example, children over the age 2 living in enrolled households are likely to have younger siblings while those of the same age in graduated households would be the oldest child.

146. To further analyse mechanisms and variables affecting the MDD of children in the graduated sample as compared to actively enrolled children, we conduct a regression analysis equivalent to section 2.8. Results of the regression are presented in Annex 11. We find that for both children in enrolled households and children in graduated households, age significantly increases the likelihood of the child meeting the MDD requirements, with slightly higher effects for graduated than for enrolled children. Thus, in absence of the programme, age of the child becomes a more important factor to receive an adequately diverse diet. For both enrolled and graduated children, the mother being married significantly increases the likelihood of children meeting MDD requirements.

147. We further find several differences in regression estimates for graduated children as compared to enrolled children. First, the mother having a formal education increases graduated children's likelihood of meeting MDD requirements, while it does not for enrolled children. This may imply that higher education is associated with a better ability (financial or cognitive) to feed children a larger variety of foods.

148. Second, graduated children living in households owning chicken are significantly more likely to meet MDD requirements. Qualitative interviews with traders revealed that many households started consuming eggs produced at home as opposed to selling them at the market as a consequence of the programme. Therefore, owning chickens may indicate those households are better able to fulfil children's consumption needs at least in terms of eggs, a major food group consumed by children in our sample.

149. Third, graduated children in households without restrictions in market access during the COVID-19 pandemic are more likely to meet MDD requirements than those graduated children who live in households that faced such restricted access.

Key findings and conclusions – Question 3.3

- While demand and willingness to buy fresh foods seems to continue after graduating from the programme, high prices of fruits and vegetables may restrict households to purchase them without the voucher.
- Examining the MDD scores for women and children in graduated households indicates that there may be a trade-off between women's and children's dietary diversity. However, as children in graduated households tend to be older than those in enrolled households, only imperfect conclusions can be drawn from this comparison.
- Both in enrolled and in graduated households age of the child seems to be indeed an important indicator of MDD. Apart from age, owning chickens and having access to a market during the COVID-19 pandemic influenced the diet diversity score of children in graduated households.

3. Conclusions and Recommendations

150. Based on the findings presented in the previous section, an overall assessment that responds to the evaluation questions is provided below. This is followed by recommendations for further action.

3.1. Overall Assessment/Conclusions

151. This evaluation covers a number of evaluation questions under the criteria relevance, effectiveness and efficiency, and sustainability. No claims of causal attribution can be made due to absence of a credible counterfactual design. Instead, we use descriptive statistics of key output and outcome indicators to gather insights into the successes and bottlenecks of the programme. All findings have to be considered as a contribution analysis, without claims of causal attribution. Further, cost-effectiveness could not be assessed due to lack of available data.

152. One further needs to keep in mind, that COVID-19 severely affected this evaluation in two ways: First, data collection was limited as no field visits were possible. Second, the voucher programme itself was operating in an unusual mode. In addition, the Tigray conflict also hampered the evaluation. One main limitation to the evaluation was the method of remote data collection for both quantitative and qualitative interviews due to the ongoing COVID-19 pandemic and the joint decision by WFP and the evaluation team to refrain from face-to-face interviewing. This method comes with substantial constraints, including the length of interviews and thus the volume of data collected, potential compromises in data reliability, and the inability to measure direct nutritional outcomes (e.g. child anthropometrics, blood samples).

153. In addition, several changes to the implementation were necessary since the end of 2019 due to the ongoing COVID-19 pandemic. Further, the outcomes of the programme need to be viewed in light of the unforeseen circumstances that may have affected beneficiaries' behaviours and access to fresh foods.

154. The findings of this evaluation might thus have been very different from what would have been found without the pandemic, which therefore also imposes limits to the transferability of the findings and conclusions to the future operations of the voucher programme.

155. The findings are discussed by the criteria Relevance, Effectiveness and Efficiency and Sustainability.

3.2. Relevance

156. Relevance of the programme was assessed with respect to the alignment with national nutrition policies, usability and appropriateness perceived by beneficiaries, and adaptation to the COVID-19 pandemic. Document review and Key Informant Interviews indicated a close alignment with strategies of national nutrition policies such as the National Nutrition Strategy (NNS)/Programme, the Seqota Declaration, and the PSNP, which acknowledge the importance of nutrition interventions for pregnant and lactating women and children under the age of two for stunting reduction in Ethiopia. The main aspects highlighted are the importance of access to nutritious fresh foods such as fruits, vegetables and eggs, demand creation and awareness-raising, as well as the strengthening of nutrient-dense foods at local markets.

157. Beneficiaries considered both components of the programme, the vouchers and the SBCC activities, as highly relevant and found the programme useful to support consumption of fresh

foods. At the same time, though, it was frequently mentioned that the voucher amounts were insufficient to purchase fresh foods for the whole family. A likely reason for this inadequacy may be that the FFV was meant to be a top-up to the PSNP transfer, but many FFV beneficiaries did not receive PSNP support. Thus, while subsidizing fresh food purchases was considered highly relevant to fulfil the needs of a nutrient-dense diet, the full potential could not be realised given the imperfect complementarity between the PSNP and FFV. In addition, food price increases, either due to fluctuating commodity prices, or potential collusion among traders, could also be a reason for the inability to purchase adequate amounts of fresh foods with the voucher.

158. The COVID-19 pandemic forced the programme to adapt very quickly to changing circumstances. The frequent face-to-face interactions during registration, market assessments, and follow-up visits were reduced since the COVID-19 outbreak. Similarly, interactive SBCC activities were suspended in February 2020. New SBCC concepts were developed and starting from July 2020, SMS and audio messages on mobile phones were disseminated. Expansion of SBCC messages through mass media outlets followed late 2020. These adaptations permitted the programme to continue reaching its beneficiaries; though it might be presumed that these remote SBCC activities might have been less tangible than the previous face-to-face interactions.

159. As a possible caveat one needs to keep in mind that the FFV programme is contingent upon access to a mobile phone. In the programme areas covered by this intervention a WFP scoping mission had indicated that nearly everyone had access to a phone within the extended family. A possible expansion of the programme to other regions should consider coverage of mobile phone access of the poorest parts of the population, though.

3.3. Effectiveness

160. For assessing if the programme worked as intended, it was analysed if the beneficiaries satisfied the eligibility criteria and if the target group was reached with the intended mix of assistance (vouchers and SBCC). Nearly every one of the enrolled households satisfied the main eligibility criteria, i.e., households with pregnant or lactating women or children below 2 years of age. However, only 20% of beneficiaries were covered by the PSNP. While the FFV programme did not target exclusively PSNP households, the intention of the voucher is to serve as a top-up to the PSNP transfer. Therefore, this may explain perceived inadequacy of the voucher amount.

161. Further, we found that almost 75% of households that recently (between October and December 2020) graduated from the programme were still eligible in terms of pregnancy status or child age. While this premature graduation may be due new pregnancies after the enrolment for a previous child, this only applies to part of the sample. Other reasons may be voluntary withdrawal or technical errors in the enrolment. However, these aspects would have to be further investigated as the data was inconclusive in this regard.

162. Delays or infrequent voucher disbursement seem to persist until the end of the expansion period, with under 60% receiving regular monthly vouchers. Results on average voucher amounts per household size indicate that the disbursement was not regularly updated by the number of household members. This may be another reason for perceived inadequacy of voucher amounts by beneficiaries. Further, irregular transfers hamper beneficiaries' and traders' possibilities to plan purchases and supply ahead and may reduce trust in the programme.

163. Despite the suspension of interactive SBCC activities due to the COVID-19 outbreak, still over half of respondents participated in coffee conversations or cooking demonstrations in 2020. A small number received messages through newly developed SBCC channels (mobile SMS and calls).

164. Further evaluation questions examined the availability of, access to and demand for fresh foods. With respect to the effects on the availability to fresh foods, it was found that a variety of products, including vitamin A rich fruits and vegetables, such as dark leafy greens and orange/dark yellow fruits and vegetables, and animal sourced products are supplied by FFV traders. While fresh foods can be procured and supplied throughout the year, seasonality plays a role in the amount and prices of fresh foods. Instable and high prices of FFV products seem to play an important role in the adequacy of voucher amounts. Strong supervision of market prices and retailer behaviour is key to ensure that fair prices are established and collusion among traders is contained in order to avoid price increases.

165. Regarding accessibility of fresh foods, it needs to be pointed out that distance to markets could be an obstacle. Access to the digital vouchers seems to be given, as beneficiaries were informed and intensively trained on how to use the mobile FFV and since further assistance was available at the market continuously throughout the implementation phase. On the other hand, general access to markets by distance depends on the means of transportation and is mixed across beneficiaries. Results indicate that especially due to long walking distances, beneficiaries may be restricted from easy access to markets. Markets were also partly affected by the COVID-19 pandemic and subsequent lockdowns, which have been lifted as of the second quarter of 2020. Around one third of beneficiaries indicated that markets could not be accessed as usual during the pandemic, including due to increased transport costs.

166. Demand for fresh food seems to be high throughout: beneficiaries make regular use of their vouchers throughout the year and several times per month. Beneficiaries mainly redeem their vouchers at fruit and vegetable traders, followed by egg traders. A majority of traders sell a large portion of their products to FFV beneficiaries. Main food groups consumed by PLW and children include grains, roots and tubers, pulses and seeds, vitamin A rich foods and other fruits and vegetables, among others.

167. Knowledge of complementary feeding as well as healthy diets during pregnancy is generally high among beneficiaries. Interestingly, although a largely consumed food group, knowledge of which foods include vitamin A is rather low. IYCF practices such as exclusive breastfeeding until 6 months and continued breastfeeding at 1 year are widely applied, however introduction of complementary foods at 6 to 8 months is less common.

168. In terms of dietary outcomes, we examined dietary diversity through various measures, keeping in mind that these are self-reported eating habit indicators elicited via telephone. No control group design for impact evaluation was possible, hence causal attribution is not possible. It was found that in terms of dietary diversity outcomes of children, the programme reached its targets for 2020 for MDD, MMF and MAD. The MDD target was reached across gender, while for MMF and MAD, the target was reached only for boys. However, compared to December 2019 levels, outcome indicators decreased by around 50 percentage points for MAD and MDD, and 7 percentage points for MMF. This may be due to a variety of reasons, including the COVID-19 pandemic and its consequences.

169. Among others, MAD seems to increase with age of the child, implying that younger children are less likely to reach dietary diversity. This may be in line with results from EQ2.4, where

complementary feeding at ages 6 to 8 months was reported as less common. As MAD further decreases with the number of children in the household, these results may indicate behavioural aspects of child feeding, i.e., that older children receive a wider variety of foods than their younger siblings. These results, however, are solely based on regression models and require further investigation.

170. The December 2020 target for minimum MDD-W was missed by 18.3 percentage points, with a decrease of 61.3 percentage points compared to December 2019. Due to the pandemic, it is difficult to indicate what may have led to this stark fall. However, we find that MDD-W increases for women who have received an FFV and were covered by SBCC, and highest for those who have attended an interactive SBCC, highlighting the importance of awareness raising.

171. The programme also aimed at increasing financial inclusion and autonomy of women as secondary aims of the programme, but the findings are inconclusive. Potential channels could be that (i) women are given the power to make food purchases for the household through the voucher and, thereby, shifting intra-household decision making; and that (ii) women gain more financial independence by making use of the mobile money wallet for other services. Qualitative evidence suggests that women (both beneficiaries and traders) gained a sense of empowerment and control due to awareness raising and training in usage. Purchase decisions seem to have been primarily made by women. Limited evidence suggests that the programme had no effect on the use of financial services by women.

3.4. Sustainability

172. There is a high willingness to continue the programme from both the governmental as well as donor side, with concrete negotiations ongoing since the end of the programme. The programme combines a high level of government engagement on federal, regional and woreda level throughout the implementation period. An expansion of the programme in additional woredas is planned and coordinated among government and other development stakeholders. There is a high willingness of the government, particularly through the Seqota Declaration, for programme continuation.

173. Government capacities have been fostered particularly in the areas nutrient gap costs and non-affordability of nutritious diet analysis in order to inform the re-adjustment of the FFV transfer value, and the contents of the SBCC implementation. Ownership structures further include private-public partnerships with the financial sector and fresh food suppliers/retailers in rural markets.

174. Another aspect of sustainability refers to the nutrition behaviour of mothers after graduation from the programme, i.e. after their children pass the second birthday. Will women continue maintaining dietary diversity for themselves and their children even when the vouchers stop? In terms of potential for sustained behavioural change after programme end, this evaluation could only give indications about the feeding and nutrition behaviours shortly after leaving the programme (while the programme itself was still ongoing). While demand and willingness to buy fresh foods seems to continue after graduating from the programme, high prices of fruits and vegetables may restrict households to purchase them without the voucher. Examining the MDD scores for women and children in graduated households indicates that there may be a trade-off between women's and children's dietary diversity. However, as children in graduated households tend to be older than those in enrolled households, only imperfect conclusions can be drawn from

this comparison. The evidence suggests that age of the child, livestock ownership and market access may play a role in the diet diversity of children after leaving the programme.

175. We find that while there seems to be high understanding of the importance of dietary diversity and willingness to continue consumption of fresh foods, financial constraints are a prohibiting factor. In terms of dietary diversity outcomes, there seems to be a trade-off between women's and children's diet. More research is needed, particularly several years after the programme ends, to investigate potential behavioural sustainability further.

3.5. Recommendations

176. Based on the findings and conclusions of this evaluation, the recommendations of the evaluation team are outlined below. It is assumed that WFP would lead a possible continuation or expansion of the FFV programme in Ethiopia, with a handover to the Government of Ethiopia at a later stage.

177. **Recommendation 1:** WFP should strengthen technical IT capacities for smooth registration, transfer and redemption process. Adaptations to reliability of mobile phone network to be considered.

178. The evaluation has revealed remaining technical issues in the pay-out of the transfers after registration. Technical problems of the IT systems should be resolved before any continuation or expansion of the programme. Although technical issues came up especially at the beginning of the programme and have been addressed since, there were still reports of registered households not receiving transfers. Particularly, issues have been reported when households change their phone numbers after registration. IT systems should permit beneficiaries to change their phone numbers quickly for continued access to the vouchers, e.g. if women gain access to their own mobile phone. (In order to avoid fraud, e.g. beneficiaries selling their SIM cards to non-beneficiaries or traders, random verification calls for identity verification may be considered.)

179. Similarly, timely and reliable monthly transfers should be ensured for optimal use of vouchers by beneficiaries. Problems with the reliability of the mobile phone network have been observed repeatedly, from which two recommendations can be derived: On the one hand, future expansions of the pilot project may be done first in areas where the mobile phone network is more reliable. On the other hand, if the programme continues to operate in regions with unreliable network, the voucher system would need to permit for more flexible solutions in order to ensure smooth transactions. One possible option could be to issue a limited number of paper-based vouchers which can be redeemed on days with recorded network problems. In order to contain operational costs of paper-based vouchers, one could consider a limit on paper-based vouchers per beneficiary, e.g. only one or two paper-based vouchers are distributed for every ten electronic vouchers used in the last year.

180. **Recommendation 2:** WFP should regularly assess the adequacy of transfer value in light of price fluctuations, and possibly of the beneficiary households' affordability.

181. While a cost of the diet analysis has proved to be an effective tool to determine an adequate transfer amount based on household sizes, a need for frequent assessments of local food prices has been identified as a consequence of increasing and fluctuating prices for fresh foods and other commodities. Given the automated voucher system one may consider an independent collection of a food price index from regional markets and an automated adjustment of the voucher amount

by the price index. (One may note that Ethiopian Public Health Institute (EPHI) was trained for such data collection.)

182. **Recommendation 3:** WFP should consider conducting an impact evaluation with control group design in order to rigorously assess causal attribution of the, so far, promising results of the pilot programme.

183. As the current evaluation was not able to measure causal attribution, WFP is recommended to carry out a rigorous impact evaluation based on a control group design. As a programme with several innovative features, this could lead to further insights of applicability in Ethiopia and other contexts. Such impact evaluation may also include the collection of anthropometric measurements of children in order to permit an analysis of stunting and wasting and possibly also blood samples of children (vitamin level, haemoglobin, anaemia etc).

184. **Recommendation 4:** Trade authorities, with WFP support, should strengthen accountability mechanisms for traders at the local markets.

185. There have been reported accounts of unfair price setting or collusion, as well as cases of unsatisfactory quality of offered products. Continuous and frequent market-level supervision plays an important role in ensuring fair market interactions. Thus, these mechanisms should be continuously supported through market surveillance and efforts to maintain accountability by traders, through spot checks or positive enforcement mechanisms (e.g. market role models). While supervision and follow-up at the market level through implementers was initially well established to ensure that quality and price complaints are addressed on the spot, disruptions of such surveillance mechanisms, as e.g. during the COVID-19 pandemic, let traders disregard standard operating guidelines. Without continuous supervision, traders in a market may collude and spontaneously increase prices when they sense increased voucher-backed purchasing power (e.g. when a group of young women enters the market, particularly when these women travelled long distances to the market and are thus price-insensitive).

186. **Recommendation 5:** WFP should focus awareness raising activities on nutrition for PLW and consider gradual phasing out.

187. While the evaluation has shown a high level of understanding of the importance of nutrient-dense foods in children's diet, low diet diversity scores indicate that women trade their own consumption of nutritious foods for that of their children. The SBCC activities should put more emphasis on stressing the need for diet diversity also for mothers. In other words, SBCC interventions need to emphasize that diet is not only important for young children, but that mother's diet diversity is also very important. The findings of this evaluation indicate that pregnant and lactating women care for their children but do not sufficiently care about themselves.

188. In addition, a gradual phase-out of the voucher amount may be considered. Currently, the voucher amount drops to zero immediately once the youngest child reaches the age of two. It appears that women's dietary scores drop quickly after graduation. Hence, instead of having the fresh food voucher amount drop to zero once children reach the age of two, a gradual reduction of the voucher amount would signal that diet diversity is still relevant even after age two. One may therefore consider maintaining the voucher eligibility for, e.g., another year with a reduced amount in order to nurture a lasting behavioural change in food habits.

189. **Recommendation 6**: WFP should investigate interactions with other WFP and government programmes (such as PSNP transfers and home gardening programmes).

190. Interactions with other programmes could reveal important insights into the effects of the FFV programme components. Further, lasting behavioural changes after graduating from the programme could be supported by other interventions that aim at self-provision of fresh foods (e.g. home gardening). Understanding how programmes work together can thus foster adjustments to FFV components.

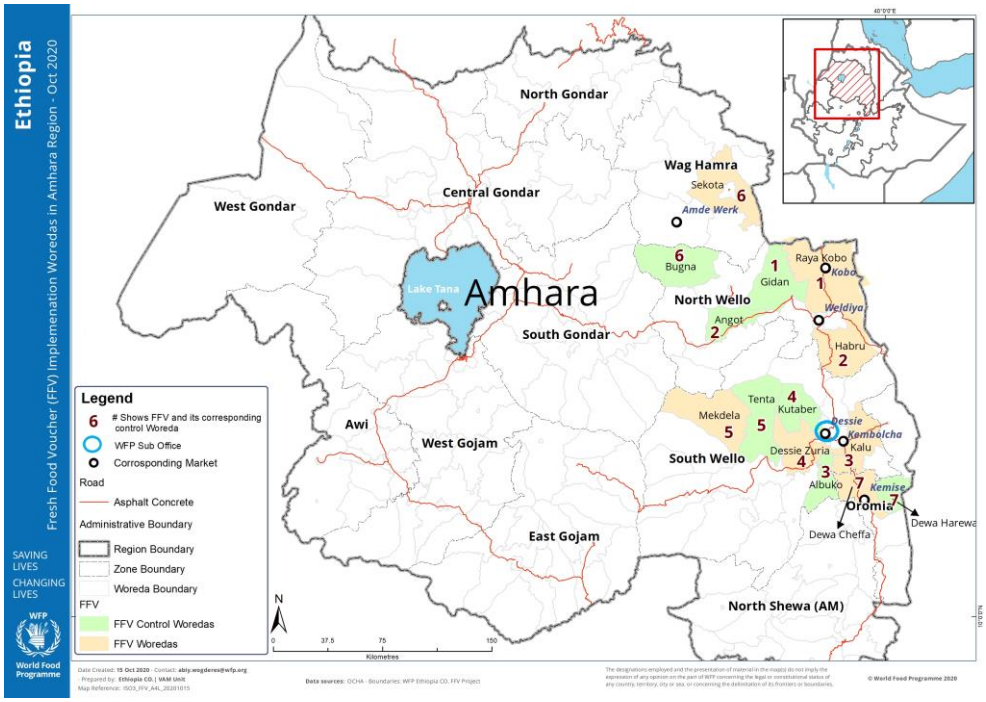
191. **Recommendation 7:** Given the importance of gender sensitive programming, WFP should continue to investigate interactions and possible synergies with other governmental and private sector actors in order to reinforce the Digital and Financial Inclusion of Women through phone-based voucher programmes.

192. The voucher scheme provides incentives for women to make regular use of mobile phones for purchase of fresh foods. This then also provides an avenue for further financial inclusion of women which has not yet been reaped. One could consider linking these women with mobile money services and microfinance providers, where the voucher scheme could also help to support mobile money savings or further (women)-group based microfinance services.

Annexes

Annex 1: Map

Figure A 1: Map of Programme Sites



Annex 2: Planned and achieved enrolment of FFV beneficiaries

Table A 1: Planned and achieved enrolment of FFV beneficiaries

Woreda	2017-2018			2019-2020				Remarks
	Target as per project doc	Enrolled at the end of Dec 2018	To be enrolled at the end of Dec 2018	Project document target	Carryover from 2019	Actively enrolled beneficiaries by 31st Aug 2020	To be enrolled by the end of the project	
	A	B	A-B	C	D	E	C-E	
All seven woredas	11,000	10,427	573	27,000	15,818	25,833 (96%)	1,167	1,680 to be discharged / graduated in the coming months

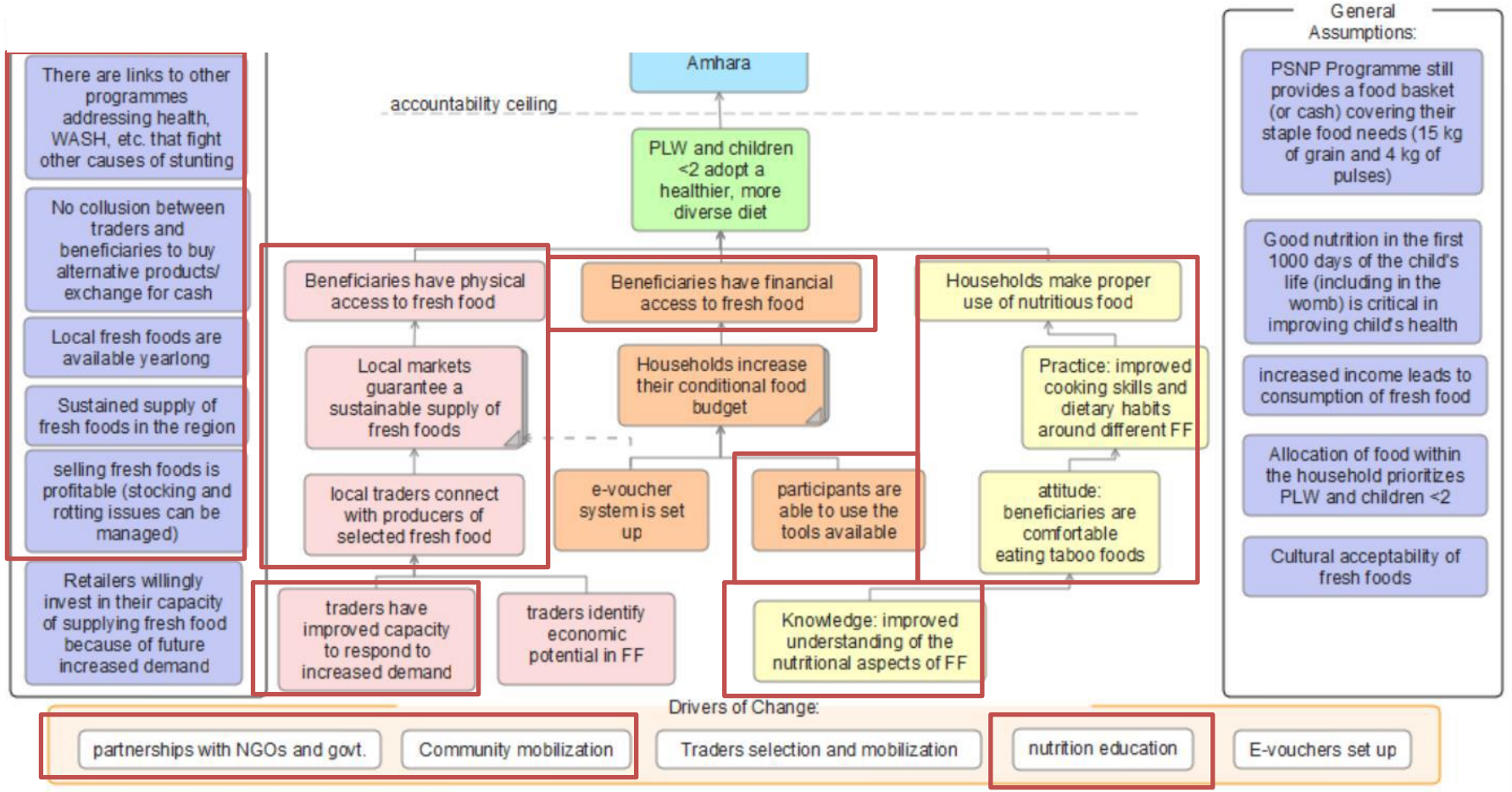
Annex 3: Documents Reviewed

Table A2: Documents gathered

Document Type
Project related documents
Impact Evaluation of WFP's Fresh Food Voucher Pilot Programme in Ethiopia
Fresh Food Voucher Programme – Design and Operational Considerations
Project Proposal to KfW – December 2018
WFP Annual Report to KfW: January – December 2019
WFP Progress Report to BMZ/KfW: January – August 2020
FFV logframe based report (December 2019)
Post Distribution Monitoring Findings PowerPoint Presentation (September 2020)
Draft SBCC Strategy September 2017
Fresh Food Voucher Retailer Engagement Baseline Report (June 2020)
Fresh Foods Value Chain & Market Assessment in Dessie Zuria Woreda, Amhara Region, Ethiopia (May 2017)
Retail training guide PowerPoint Presentation
Ethiopian Policy Documents
Food and Nutrition Policy (November 2018)
National Nutritional Program 2016-2020
Seqota Declaration Implementation Plan (2016-2030)
The Seqota Declaration: Innovation Phase Investment Plan 2017-2020
Other documents
Cost of the Diet – FFV End-Line Evaluation
Evaluations reviewed
Indicators included mother's diet; mother's anthropometry; mother's workload and time use; maternal infant and young child feeding knowledge, attitudes, and practices; child anthropometry; and hygiene conditions in the home
Berhane et al. (2020). Evaluation of the Nutrition-sensitive Features of the Fourth Phase of Ethiopia's Productive Safety Net Programme. IFPRI Strategy Support Program | Working Paper 140 http://ebrary.ifpri.org/utils/getfile/collection/p15738coll2/id/133685/filename/133897.pdf
Abbot, Daniel &, Fekau, Habtamu. (2020). Growth through Nutrition Midterm Evaluation Program Review Report https://resourcecentre.savethechildren.net/node/17535/pdf/Growth%20through%20Nutrition%20Mid term%20Evaluation%20Report.PDF
Action Contre la Faim (ACF). 2012. Meta-evaluation of ACF Fresh Food Voucher Programmes https://www.actioncontrelafaim.org/wp-content/uploads/2018/01/acf_fresh_food_voucher_meta_evaluation_2012.pdf
Hirvonen et al. (2019). Decentralized Evaluation: Impact Evaluation of WFP's Fresh Food Voucher Pilot Programme in Ethiopia. 10/2017-1/2019. Evaluation Report. https://docs.wfp.org/api/documents/WFP-0000106838/download/?_ga=2.137059760.165364307.1610535203-1455894063.1605707375 accessed 01/14/2021

Annex 4: Theory of Change

Figure A2: Theory of Change



Note: ToC elements included in the Evaluation Matrix framed in bold.

Annex 5: Evaluation Matrix

Table A3: Evaluation Matrix

EQ No.	Evaluation questions (EQs)	SEQ No.	Sub-Evaluation questions (SEQs)	Method	Quantitative Indicators/ Qualitative Guiding Questions	Tool	Data reliability
1. RELEVANCE							
1	Is the project aligned with the national government's policies and strategies to support the reduction of stunting via increased diet diversity?	1	How do the WFP programme activities respond to national government's policies and strategies to support the reduction of stunting via increased diet diversity?	Qualitative	What are the national policies and guidelines regarding reduction of stunting? What are the main strategies recommended in national policies and programmes to reduce stunting via increased diet? How were these recommendations included in the programme design?	KIIs/document review	2=fair
2	How relevant has the approach been to beneficiaries?	1	Is household decision making regarding food purchases and nutrition aligned with handling the mobile money vouchers?	Qualitative	How are the decision related to food purchase and nutrition taken within the household? Who in the household has access to the mobile money vouchers?	IDIs	
		2	Do the beneficiaries consider the FFV and SBCC useful?	Qualitative	Are FFV easy to use and promoted practices easy to adopt? What are the barriers to use FFVs and adapt nutrition behaviours? Is the amount of the voucher appropriate?	IDIs	
		3	How useful was the programme for beneficiaries in light of the COVID-19 pandemic?	Qualitative	Did beneficiaries use the vouchers more or less during the pandemic? Why or why not?	IDIs	
		4	How relevant was the programme for beneficiaries with male/female children?	Qualitative	Are there differences in usage of the FFV based on gender of children? Why?	IDIs	
3	How did the programme interventions adapt to the COVID-19 pandemic?	1	What adjustments were made to the implementation due to the COVID-19 pandemic?	Qualitative	Changes to implementation due to COVID	KIIs	

2. EFFECTIVENESS AND EFFICIENCY							
1	Did the project reach the intended beneficiaries with the right mix of assistance?	1	Do households receiving the programme meet the eligibility criteria (PLW/child 6-23m, PSNP)?	Quantitative	% of beneficiary households who are eligible	Household Survey	2=fair
		2	Did beneficiary households receive the monthly transfer as intended?	Quantitative	**% of beneficiary households receiving the correct voucher amount based on their household size**	Household Survey, M&E	
					% of beneficiary households receiving a regular transfer on a monthly basis	Household Survey, M&E	
		3	Were beneficiaries aware of SBCC activities?	Quantitative	% of beneficiaries indicating awareness of SBCC activities (cooking demonstrations, coffee conversations, radio shows, information material)	Household Survey	
		4	Did beneficiary households participate in SBCC activities?	Quantitative	% of beneficiaries attending SBCC activities (cooking demonstrations, coffee conversations)	Household Survey	
		5	Did beneficiary households receive SBCC messages?	Quantitative	% of beneficiaries recalling messages from (1) cooking demonstrations, (2) coffee conversations, (3) radio shows	Household Survey	
		6	Did beneficiaries receive information on healthy breastfeeding and child feeding practices?	Quantitative	% of beneficiaries who received visit from HEW/WDA/HDA and recall topics covered	Household Survey	
2	What are the effects of the project on availability of fresh foods?	1	Can traders provide a variety of fresh foods all year?	Quantitative	**% of traders that offer at least one of the following categories: (1) dark leafy greens, (2) orange vegetables, (3) orange fruit, (4) animal-sourced food**	Trader Survey	
				Quantitative	% of traders who can supply fresh foods all year	Trader Survey	
				Quantitative	% of traders who understand how to procure fresh food	Trader Survey	
				Quantitative	**% of traders who understand how to handle fresh food**	Trader Survey	
				Qualitative	**Bottlenecks to supply of fresh foods**	IDIs	
		2	How did prices of fresh foods change during the program period?	Quantitative	% of traders offering products with a stable price level	Trader Survey	
				Quantitative	% of traders who know how to set fair and competitive prices	Trader Survey	
				Qualitative	Reasons for price fluctuations	IDIs	

		3	Are beneficiaries satisfied with the quality of fresh food available at participating traders?	Quantitative	**% of beneficiaries satisfied with quality of fresh food available at participating traders** / Level of satisfaction with quality of fresh food available at participating traders	Household Survey	
				Qualitative	What are necessary improvements to quality of offered food to increase demand?	IDIs	
3	What are the effects of the project on access to fresh foods?	1	Do targeted beneficiaries know how to use mobile money voucher services?	Quantitative	**% of targeted beneficiaries who know/were trained on how to use mobile money voucher services**	Household Survey	
		2	Do beneficiaries have easy access to markets where participating traders are selling fresh foods?	Quantitative	Minutes from nearest market where they can redeem the FFVs	Household Survey	
		3	Did access to markets and fresh food change during the COVID-19 pandemic?	Quantitative	Access to market during COVID-19 pandemic	Household Survey	
4	What are the effects of the project on the demand for fresh foods?	1	Are beneficiaries making use of the mobile money vouchers?	Quantitative	**% of beneficiaries performing mobile money voucher transactions**	M&E	
					frequency of FFV usage	Household Survey, M&E	
					% of voucher value redeemed on fresh food	M&E	
					% of fresh food sold to FFV beneficiaries	Trader Survey	
		2	What is beneficiary's knowledge about favourable IYCF?	Quantitative	% of beneficiaries who know about beneficial infant and child feeding practices (number of correct answers)	Household Survey	
					% of beneficiaries who know about beneficial nutrition for pregnant and lactating women	Household Survey	
		3	How do beneficiaries practice IYCF?	Quantitative	% of beneficiaries practicing beneficial infant and young child feeding practices	Household Survey	
5	Did the intervention produce the expected nutritional results?	1	Has the target on children's minimum acceptable diet been reached?	Quantitative	**Minimum Acceptable Diet (MAD) scores for children (6-23 months) (%)**	Household Survey	
		2	Has the target on mothers' dietary diversity been reached?	Quantitative	**Minimum Diet Diversity for women (MDD-W) (15-49 years) (%)**	Household Survey	
		3	Has the target on children's dietary diversity been reached?	Quantitative	**Minimum Diet Diversity (MDD) for children (6-23 months)**	Household Survey	

		4	Has the target on children's dietary diversity been reached across gender?	Quantitative	**Minimum Diet Diversity (MDD) for children (6-23 months)**	Household Survey	
6	Is there an increase of financial inclusion and financial autonomy among female beneficiaries?	1	Do female beneficiaries use more the mobile wallets for transactions other than before the intervention?	Quantitative	% of women using mobile money wallets for other transactions Frequency of usage mobile money wallets for other transactions	Household Survey	
		2	Do female beneficiaries perceive themselves as more autonomous regarding financial decisions than before the programme?	Qualitative	Perceived financial autonomy and inclusion of beneficiary mothers	IDIs	
7	Was the intervention efficient compared to possible alternatives?	1	Why was this approach considered the most efficient?	Qualitative	Which approaches were considered and how was the approach chosen? Compared to alternatives, was the chosen approach the most efficient?	KIIs/document review	
4. SUSTAINABILITY							
1	What is the government readiness to take over the programme?	1	What is the government readiness to take over the programme?	Qualitative	What government engagement strategies were applied throughout the programme? What capacities are still needed? What needs remain for a full handover within the PSNP-V? How does the FFV programme currently align with other nutrition policies, notably the Seqota declaration and NNP?	KIIs	2=fair
		2	What are the anticipated ownership structures?	Qualitative	How are public-private partnerships considered for the continuation of the programme? How is community participation and ownership incorporated?	KIIs	
2	Is the programme capable of overcoming future challenges and bottlenecks?	1	How ready is the programme to deal with potential challenges and bottlenecks?	Qualitative	What were the challenges in implementing the vouchers and SBCC activities over time? What are the strategies to overcome these in the future?	KIIs	
3	Will behavioural changes related to nutrition last after the programme?	1	Is the programme sustainable in terms of nutritional status of graduated beneficiaries?	Quantitative	Nutrition outcomes (MAD, MDDW, MDD, MMF) for graduated beneficiaries	Household Survey	

Note: Indicators from logic model highlighted in bold font.

Annex 6: Full Methodology

The evaluation of the FFV programme serves the purposes of organizational learning and accountability on two levels. On the one hand, it assesses the implementation of the activities. On the other hand, it investigates the reception of the intervention among the targeted population, that is, the fresh food traders and the beneficiary mothers and respective children. One needs to keep in mind, however, that the accountability aspect is severely restricted by the COVID-19 pandemic: on the one hand, because of its impact on the beneficiaries, in that outcomes of beneficiaries would have been different without the pandemic; and, on the other hand, due to the limitations imposed on the data collection tools. In the following subsection, we present the proposed approach and methodology, describing in detail the evaluation matrix, the tools and the sampling strategy.

1. Approach and Methodology

We proposed a mixed-method approach including both quantitative and qualitative analyses. The quantitative analysis consisted of a detailed description of outputs and outcomes relating the programme beneficiaries and traders. We made use of data collected through phone surveys, which we cross-checked with M&E data shared by WFP (see sub-section 3 for more details on the M&E data). During the inception phase, besides policy-related documents, several programme background documents and data were gathered and thoroughly reviewed (see Annex 3). Programme documents such as project proposals, project narrative, and annual reports were used as sources to inform the evaluation design and evaluation questions.

The qualitative analysis consisted of KII with stakeholders and IDI with beneficiaries in order to provide a better insight into the quantitative results and included multiple perspectives and opinions. Further, the use of multiple data sources and methods ensured the coherence and the impartiality of the results.

The evaluation matrix (Annex 5) was informed by the ToC (Annex 4) and linked the overarching DAC criteria set by the ToR to respective high-level EQs. These EQs are broken down into more specific sub-EQs and further translated into indicators. The evaluation matrix also indicates which questions were answered through which tools.

We aimed to evaluate the programme through the lens of the DAC criteria of Relevance, Effectiveness and Efficiency, and Sustainability.[29] The original evaluation matrix further included Cost-Effectiveness, which was dropped at a later stage due to lack of available data.

In the following, we describe in more detail the proposed scope of the evaluation along each DAC criterion.

1.1. Relevance

Evaluation questions associated to the Relevance criterion explored the extent to which the intervention objectives and design respond to national policies and strategies, as well as

[29] For more details, see http://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm.

beneficiaries' needs in an adapted manner. For the latter, we employed qualitative methods to discover the alignment of the vouchers with household decision making, its usefulness and practicality and any differences regarding gender of children.

1.2. Effectiveness and Efficiency

Effectiveness refers to the ability of the programme to reach its goals in an economic and timely way. To assess this, we first assessed whether the programme had (1) reached out to the intended population and (2) implemented the planned activities. Furthermore, we measured the extent to which beneficiaries had (3) adopted and (4) understood the services provided by the programme as well as (5) the benefits of a diverse diet and adjust their feeding practices accordingly. This allowed us to assess whether the (6) intended nutritional goals have been reached.

Secondly, effectiveness was assessed by investigating the provision of fresh foods among traders associated to the programme. On the one hand, this was done by measuring supply at the product level, that is, whether (1) traders offer fresh foods and whether (2) they can supply such foods all year around. We also assessed the knowledge of traders to (3) procure and (4) handle fresh food. On the other hand, supply was assessed at the price level by gathering information on (5) the stability of prices throughout the year and on (6) the price-setting approaches. Finally, we measured the use of FFVs by traders by estimating (7) the share of the fresh foods sold to FFV beneficiaries.

1.2.1. Key output indicators

To measure the outputs along the ToC as outlined above, we presented several output indicators related to traders, beneficiaries' usage of the FFV and the SBCC activities. Key output indicators as identified by the logframe are presented by highlighting end-line levels against those from baseline (2018 pilot), 2019 PDM data, and the targets for 2020. However, due to unclarity of the construction of several output indicators in the logframe, the values measured at end-line may only be of limited comparability to the target and earlier levels. Table A4 presents the original logframe indicators. Yellow and green highlighted indicators are included in the evaluation matrix. Yellow indicates an original indicator, green an added indicator. Italic indicates the measurement is different e.g., due to missing information for replicability.

Table A4: Output Indicators related to Beneficiaries and Traders

Output Indicator (Beneficiaries)	**Baseline 2018**	**2019**		**Target 2020**
		August	**December**	
Beneficiaries satisfied with quality of fresh food available at participating traders (%)	*n.a.*	*n.a.*	*n.a.*	***≥80***
USD value transferred to beneficiaries though mobile-based vouchers	n.a.	1,236,341	1,975,001	5,000,000
Targeted beneficiaries who know how to use mobile money voucher services (%)	*n.a.*	*n.a.*	*n.a.*	***100***
Targeted beneficiaries performing mobile money voucher transactions (%)	n.a.	98	94	**100**
Voucher value redeemed on fresh food (i.e., value redeemed/valued transferred) (%)	n.a.	80	80	**100**
Number of beneficiaries redeeming mobile money transactions from the traders	n.a.	13,152	16,935	**27,000**

Number of targeted beneficiaries with mobile money wallets	n.a.	15,410	21,653	**27,000**

Output Indicator (Traders)	**Baseline 2018**	**2019**		**Target 2020**
		August	December	
Type of and quantity of nutritious fresh food has increased per selected merchant (%)	*n.a.*	*71*	*71*	***>80***
Licensed traders participating in FFV (%)	100	100	100	**100**
Number of traders who know how to analyze their financial transactions and stock levels	n.a.	34	n.a.	**90**
Traders who understand WFP standard operating procedures (SOPs) (%)	n.a.	100	100	**100**
Number of traders who understand how to procure and handle fresh food	*n.a.*	*44*	*n.a.*	***≥75***
Numbers of traders operating under WFP SOPs	10	220	242	**300**
Traders know how to use mobile money voucher system (%)	n.a.	100	n.a.	**100**
Number of targeted traders with mobile money wallets	10	220	242	**300**
Added: % of traders who can supply fresh foods all year	n.a	n.a.	n.a.	n.a.
Added: % of traders offering products with a stable price level	n.a	n.a.	n.a.	n.a.
Added: % of traders who know how to set fair and competitive prices	n.a	n.a.	n.a.	n.a.

Notes: Yellow and green highlighted indicators are included in the evaluation matrix. Yellow indicates an original indicator, green an added indicator. Italic indicates the measurement is different e.g., due to missing information for replicability.

Output indicators were both captured through the end-line surveys with beneficiaries and traders as well as from the M&E systems, notably the HelloCash Analytics. For these, data was shared with the evaluation team for the period of June 2018 to September 2020 which can be used for cross validating the end-line data collected in December 2020.

1.2.2. Key outcome indicators

For outcomes, we limited the indicators to the child and PLW nutrition outcomes while omitting household level indicators. This was decided during inception due to the limited capacity of the phone survey and prioritization of outcomes of interest. Table A5 shows the original outcome indicators from the logframe, where yellow highlighted indicators are also included in the evaluation matrix.

Table A5: Outcome Indicators related to Beneficiaries

Outcome indicator	**Baseline 2018**	**2019**		**Target 2020**
		August	**December**	
Minimum Acceptable Diet Scores (**MAD**) for children (6 to 23 months) (%)	22.3	40.2	93.3	**42.3**
Minimum Diet Diversity for Women (**MDDW**) for women (15-49 years) (%)	3.0	70.0	93.0	**50.0**
Minimum Diet Diversity (**MDD**) for children (6 to 23 months) (%)	22.3	59.6	96.6	**42.0**
Minimum Meal Frequency (**MMF**) for children (6 to 23 months) (%)	86.8	63.2	96.6	**90.0**
Household Diet Diversity Score (HHDDS)	6.8	8.4	9.52	**>5**
Food Consumption Score (FCS) - % of HHs which never consumed: Protein	7.0	8.0	0.0	**< 5**
Vit A rich food	66.7	5.0	0.0	**<5**
Hem iron rich food	89.0	98.0	50.0	**< 50**

For child nutrition outcomes, we focussed on the MAD, MDD, and MFF. For women, we assess the MDD-W. These indicators were constructed following the 2017-2021 Programme Indicator Compendium shared by WFP:

To calculate the MAD, we first defined 7 food groups according to World Health Organisation (WHO) guidelines, including grains, roots and tubers (1), legumes and nuts (2), dairy products (3), flesh foods (4), eggs (5), vitamin-A rich fruits and vegetables (6) and other fruits and vegetables (7). We then defined a dichotomous variable for the MAD, which is =1 if the child has eaten at least four of the food groups, and 0 otherwise. To measure the MMF, we defined a second dichotomous variable which is =1 if a child of 6-8 months is breastfed and gets food at least two times per day (1), if a child of 9-23 months is breastfed and gets food at least three times per day (2) and if a child of 6-23 months that is not breastfed gets food at least four times per day (3). The MAD is defined as a dichotomous variable that is =1 if a child of 6-23 months meets the requirements of both MDD and MMF.

Similar to the key outputs, we assessed the levels of nutrition outcomes at end-line (i.e. December 2020) with target levels from the logframe[30].

We further employed several approaches to analyse the nutrition outcomes in more detail:

1) Summary statistics (by various subgroups):

We first presented the MAD and its sub-components (MMF and MDD) by gender and for children of different age groups in our sample. While MDD, MMF and MAD scores are typically measured for children of ages 6 to 23 months, we presented the MDD for all children up to the age of four years as we collected the data for these children. In the remaining analysis, we restricted the sample to children 6-23 months.

We examined the nutrition outcomes for children older than 24 months for two types of families: Those families where the youngest child is younger than 24 months, such that the household is still eligible for FFV, and those families where the youngest child is older than 24 months and the household therefore not receiving FFV anymore. This comparison provided an indication about the sustainability of behavioural change after the voucher support stops.

For MDD-W, we presented the results for the age group of 15 to 49 years.

The subgroup analysis covers the following sample splits:

- By receipt of an FFV in the past 6 months
- By exposure to SBCC components

2) Regression analysis:

i. Nutrition outcomes for children:

[30] One should note again that any differences between actual and target values could also be due to the COVID-19 pandemic.

For the MAD, MDD and MMF we run the following regression model:

$$outcome_{chr} = \beta_{0+} + \beta_1 season\_birth_{chr} + \beta_2 regyear_{hr} + \beta_3 season\_birth_{chr} * regyear_{hr} + \beta_4 age_{chr} + \beta_5 age_{chr}^2 + \beta_6 age_{chr}^3 + \beta_7 woreda_r^3 + \varepsilon_{chr} \quad (1)$$

Where:

- the unit of analysis is child c aged 6-48 months, living in household h located in woreda r.
- $season\_birth$ indicates season of birth for child (lean/harvest).
- $regyear$ indicates year of registration (2018/2019).
- age, age^2, age^3 is a cubic polynomial of child's age.
- $woreda$ dummies indicating woreda.
- Standard errors are clustered at the household level.

Note that the presented model is the basic specification. We can then further add the following controls: gender of child, child is breastfed (yes/no), household size, married (yes/no), mother has formal education (/yes/no), number of children, child is first born (yes/no), household has electricity (yes/no), owned livestock type dummies, market distance, market access during covid (yes/no) PSNP support (yes/no).

ii. Mothers' nutrition outcomes:

For the regression analysis of MDD-W for PLW, we run a similar regression:

$$outcome_{hr} = \beta_{0+} + \beta_1 regyear_{hr} + \beta_2 children623 + \beta_3 children24^{+} + \beta_4 woreda_{wr}^3 + \varepsilon_{hr} \quad (2)$$

Where:

- the unit of analysis is woman aged 15-49 in household h, located in woreda r.
- $regyear$ indicates the year of registration (2018/2019)
- age category dummies for having any children aged 6-23,24-48. The omitted category is children aged 0-6 months
- $woreda$, dummies indicating woreda
- Standard errors are clustered at the household level

Again, the presented model is the basic specification, and further control variables can be added: mother is currently breastfeeding (yes/no), household size, married, mother has formal education(yes/no), number of children, household has electricity (yes/no), owned livestock type dummies, market distance, market access during covid (yes/no), PSNP support (yes/no).

1.3. Sustainability

Sustainability is measured by the extent to which the activities and intended effects of the intervention are likely to continue. To do so, we used KII to comprehend the political, financial and institutional viability of the programme. One aspect that was investigated as well as how the programme may have shaped national policies. Additional KII provided insights on the challenges and bottlenecks that the programme faced, and the strategies used to overcome them. Finally, we collected data from graduated beneficiaries about their nutrition practices in order to inform on whether behavioural changes last after the programme.

We also examined sustainability at the individual level by "sustained behavioural change" in that we examined the nutrition behaviour of mothers after they graduated out of the program: Do they maintain different nutrition patterns even once they are not eligible for food vouchers anymore?

To answer this question, we first examined the MDD for both PLW and children in graduated households. We further looked at potential drivers by comparing regression results for children in enrolled and graduated households using the equation (1) above.

1.4. Cost-Effectiveness

A cost-effectiveness analysis was planned by measuring (1) the costs of the implementation of WFP activities per beneficiary as well as (2) the FFV transfer cost per beneficiary. However, access to relevant internal documentation could not be granted to the evaluation team by the time of reporting. Therefore, no evidence on cost-effectiveness of the programme could be performed for this report.

2. Site Mapping and Sampling

The evaluation sample covers six of the woredas where the programme was scaled up, i.e., Habru, Raya Kobo, Dessie Zuria, Dawa-Cheffa, Mekdela and Kalu. The seventh woreda, Sekota, was dropped from the sample because of the ongoing outage of the phone network connectivity in the area due to a conflict in the neighbouring Tigray Region. The data collection was entirely conducted remotely over the phone; hence no sites were visited. In this section we present the sampling strategy for the quantitative and qualitative analyses.

As regards the quantitative analysis, we sampled from beneficiaries and traders. Sample size for each group was chosen to maximize the targeted population covered given the available budget for the evaluation.

For the beneficiary group, a sample size of 1380 was chosen to be drawn from the entire beneficiary population across the six woredas, as per the scope of the evaluation. The sample was homogeneously distributed across the woredas, i.e., approximately 230 households who registered during 2018 and 2019 were targeted for an interview in each woreda.

The sample was stratified by the type of vulnerability category targeted by the programme: pregnant or lactating mothers of children between 6 and 23 months at the time of registration. In addition, we also differentiated between two types of households based on voucher redemption: those households that redeemed no or few vouchers and those that redeemed vouchers with higher frequency. For this reason, we oversampled the former category by making sure to include up to 50 households that redeemed from zero up to three vouchers for each woreda-stratum combination in the sample. Based on the population share of each of these categories, adequate sampling weights were generated and used for the regression analysis to ensure representativeness. All households with children up to four years of age in the sample described above were interviewed.

In case the originally sampled household was not available for interview or refused the interview, the household was replaced by a household with similar stratification characteristics. This was the case for around 12% of the sample.

Table A6: Household Characteristics by Woreda

Woreda	Total Number of HHs	% Pregnant at Registration	% lactating at Registration	% Low FFV Redemption
Dawa-Cheffa	230	50.0%	50.0%	25.7%
Dessie-Zuria	233	49.8%	50.2%	8.6%
Habru	228	49.1%	50.9%	21.5%
Kalu	230	50.4%	49.6%	39.6%
Kobo	229	50.2%	49.8%	24.0%
Mekdella	230	50.0%	50.0%	33.5%
Total	1380	49.9%	50.1%	25.5%

Table A6 presents summary statistics for household characteristics used for stratification by woreda. Note that the analysis is not disaggregated by woreda and the sample is not representative of the beneficiary population per woreda. Rather, as described above, the sample is representative of the beneficiary population by pregnancy status and frequency of voucher redemption across all woredas.

Table A7: Household Eligibility Characteristics

	count	%
Household Size		
up to two HH members	21	1.5
3 to 5 HH members	678	49.1
6 and more HH members	681	49.3
Ages of children in household		
0-5 months	95	6.9
6-23 months	903	66
24-48 months	642	46.9
above 48 months	3	0.2
PSNP Households	270	19.6

The average household had 5.6 members, with around half of the households having three to five household members and another half six and more members. As shown in Table A7, only 1.5% of households had up to two members. The majority of households, 66%, had children between the ages of six to 23 months living in the household. Another large share, almost 47%, had children between 24 and 48 months. Almost 20% of the respondents indicated their household had received a PSNP transfer in the last 12 months.

In total, we received data from 1,690 children out of the 1,380 households (Table A8). Equal shares of 16.8% of children were of the ages six to eleven months and 12 to 17 months, respectively. Another 20.9% were between 18 and 23 months. The gender of children in the sample is balanced, with around 50% male and female children.

Table A8: Child Characteristics

	count	%
Age group		
below 6 months	96	5.7
6 to 11 months	284	16.8
12 to 17 months	284	16.8
18 to 23 months	353	20.9
24 to 48 months	670	39.6
above 48 months	3	0.2
Gender		
Male	854	50.5
Female	835	49.4
Refuse to answer	1	0.1
Total	1690	

As for the trader survey, we interviewed 140 of them, i.e. more than 72% of the HelloCash registered traders in the programme across the six woredas. Traders in the sample were also be sampled from each of six woredas in the evaluation sample. Table A9 presents number and percent of traders by woreda and type of product sold.

Table A9: Trader Characteristics

	count	%
Woreda		
Dawa Cheffa	12	8.6
Dessie Zuria	37	26.4
Habru	18	12.9
Kalu	11	7.9
Mekdela	30	21.4
Raya Kobo	32	22.9
Product Type		
Fresh Vegetables or Fruits	123	87.9
Animal Sourced Products	16	11.4
Total	140	

Qualitative sample

The qualitative analysis includes beneficiaries and traders through IDI and programme implementation stakeholders through KII. KII respondents include high-level representatives of the programme as well as local stakeholders including programme coordinators, as well as nutrition, agriculture, and trade officials on woreda level. In total, 13 KII were conducted.

Table A10: Key Informants interviewed

Category	Count	Sex	Responsibility
Government stakeholders	1	M=1 F=0	• Seqota declaration senior officer (1)
Central level KII (WFP country office staff)	5	M=2 F=3	• Nutrition Lead at the WFP ETHCO office (1) • Stunting reduction coordinator (1) • Nutrition consultant/Analyst (1) • FFV manager at ETHCO (1) • Marketing manager at ETHCO (1)
Woreda level KII	8	M=6 F=2	• Woreda FFV coordinator (3) • Woreda Agriculture (3) • Woreda Health centre (1) • Woreda Trade office (1)

Beneficiaries for the IDIs were sampled from the database of beneficiaries not participating in the household survey in order not to bias responses to either data collection as the two occurred in parallel.

Table A11: IDI Beneficiary Characteristics

Characteristic		Count
Age	18-25	7
	26-35	6
	36-45	2
Sex	male	0
	female	15
Marital status	married	14
	Single	1
	divorced	0
Family size	≤ 2	0
	3-5	8
	≥6	7
Woreda	Kobo	4
	Dessie Zuria	3
	Habru	2
	Kalu	2
	Dawa Chawa	2
	Mekdela	2

Characteristics of beneficiary IDI respondents are presented in Table A11. Respondents were between 18 and 45 years old, female, and from the six sampled woredas.

For trader IDI, we interviewed both FFV traders and non-participating traders[31]. Respondents traded in fruits, vegetables, and eggs were between 18 and 35 years old. Due to the lack of available respondents in the list of traders, the majority of respondents were based in Seqota, Kobo, and Dessie Zuria.

[31] These non-FFV traders were originally in the list of FFV traders shared by WFP, indicating that they had been registered as FFV traders but did not end up participating in the programme.

Table A12: IDI Trader Characteristics

Characteristics		FFV	Non-FFV
Food items sold	Fruits	7	8
	Vegetables	7	8
	Eggs	2	2
Age	18-25	2	2
	26-35	3	4
	36-45	3	0
Sex	male	3	0
	female	7	6
Marital status	married	7	7
	Single	1	1
	divorced	0	0
Family size	≤ 2	1	0
	3-5	8	3
	≥6	2	2
Woreda	Seqota	1	2
	Kobo	6	1
	Dessie zuria	1	2
	Habru	0	0
	Kalu	0	0
	Dawa Chawa	1	0
	Mekdela	0	1
Total		**10**	**6**

3. Data collection methods and tools

As can be seen in the evaluation matrix, we combined different tools and methods to answer each of the assessment criteria and EQs. The evaluation used household and trader surveys as main data sources (where survey data could only be collected via phone because of the COVID-19 pandemic). However, for the EQs for which information was available from secondary sources provided by WFP, data was validated or complemented as appropriate.[32] Multiple secondary data sources were combined. Data on beneficiaries and traders were retrieved from HelloCash analytics[33] until September 2020 as well as from the WFP M&E system. The M&E data shared by WFP from came from three databases: i) the SCOPE beneficiary database[34] covering eligibility characteristics of the households registered in the programme; ii) programme enrolment lists[35] updated as of October 2020; iii) dashboard data including beneficiaries' transaction details for the period September to November 2020.

[32] Given the obvious limit to the level of depth that can be covered by the survey, the validation will not be possible for details such as historical voucher number and value and use of voucher.

[33] For beneficiary households, the dataset includes household identification, total amounts and values of transactions the household received from WFP (including dates of transactions), amount and value of redemptions, balance in the Hello Cash account, refunded amount to WFP of unused amounts, total number and value of fresh food transactions from beneficiaries to merchants per type of business. As regards traders, the dataset includes: identification information, market name, type of business, total number and value of fresh food redemptions, balance in the HelloCash account, total number and value of withdrawals from the HelloCash account.

[34] The SCOPE WFP beneficiary database includes detailed information of the beneficiary households, collected at the time of registration. Relevant information includes household and personal identification, kebele, relation to the household head, gender, date of birth, household size, breastfeeding status, pregnancy status.

[35] The WFP enrollment dataset includes: household identification, household size, beneficiary status.

Due to the COVID-19 pandemic data collection was only possible via remote tools, i.e., telephone interviews, due to WFP's ethical requirements that no face-to-face contact was permitted. For these reasons, data reliability, quality and the depth of information and level of detail is clearly lower than with fully-fledged face-to-face household surveys. The requested timing of the evaluation, however, did not permit any other alternative.

Quantitative tools used to collect primary data are surveys with (i) beneficiary households and (ii) participating FFV traders. The primary respondents for beneficiary interviews were women of reproductive age. The household survey focused on nutrition outcomes as measured by indicators outlined above, in particular related to women's and children's nutrition. Further, knowledge, attitudes, and practices (KAP) were captured through several questions asking about breastfeeding and young child nutrition. Another module includes a short household background section. Finally, the questionnaire captured awareness of, participation in and learnings from SBCC activities.

The trader survey captures information about the type of business and products the retailer supplies and sells, their practices of procuring, handling and marketing fresh foods and their engagement in the FFV programme.

Qualitative tools consisted of semi-structured individual interviews. We first conducted a set of KII with (i) government stakeholders and (ii) project implementers prior to the quantitative data collection. The KII with the project implementers was aimed at better understanding the rollout of the programme including the achievements and challenges faced. Further interviews with relevant stakeholders provided a broader perspective on the FFV programme in regard with national policies[36] and its long-term viability (i.e. sustainability).

In addition, and parallel to the quantitative data collection, we conducted IDI with beneficiary mothers to collect, among other information, feedback from the services received, challenges faced, expectations and perceived changes (particularly concerning financial autonomy). Further, access to phones and usability of the digital voucher system were covered in the IDIs to shed light on the relevance of the intervention for vulnerable groups.

IDI with traders elicited information about the bottlenecks to access and supply of fresh foods, reasons for price fluctuations, and their experiences with the mobile voucher programme. Qualitative data from IDI also informed the quantitative results from surveys by investigating more details into potential pathways and bottlenecks of fresh food provision.

Whenever possible, multiple sources of evidence were used for triangulation. In particular, quantitative data from phone surveys were compared against M&E data from the programme as well as from qualitative evidence gathered via IDI and KII. For example, certain information collected from household interviews, such as reception and transfer values as well as eligibility criteria, was cross-checked with data from M&E systems. The qualitative data from IDIs gave more in-depth insights into experiences with the FFV programme, complementing the aggregated results on usage of the vouchers and the demand and supply of fresh foods from the survey data. Potential gaps of evidence that may lead to inconclusive findings was clearly indicated.

[36] Such as the NNP, FNP, and the Seqota Declaration implementation plan.

4. Limitations and Risks

The COVID-19 pandemic and limitations related to remote data collection

An undeniable limitation lies in the data collection process. Given the COVID-19 pandemic situation in Ethiopia, WFP has decided that face-to-face interviews were not feasible, neither for the surveys nor for the IDI or KII. Instead, we conducted phone interviews. This limited the observation of behaviour and body language, as well as any visual aids to assist in the interviewee. This is typically particularly relevant for qualitative interviews but may also have limited the enumerator's ability to react to non-verbal cues of the respondent's understanding or confidence in answering the question. Meta data on the enumerator's impression of the respondent's engagement in the interview was collected, however the reliability of this information is also difficult to judge and only allows for a limited ex-post evaluation of the accuracy. Further, while interview audits are possible through spot recording of phone calls, the evaluation team refrained from using this tool as recordings potentially cause discomfort or unwillingness to participate in the interview.

A very important limitation is also the inability to collect anthropometric measurements. Also, due to practical reasons, phone interviews limit the lengths of questionnaire and therefore the quantity of data collected.

Regarding the phone questionnaire, nutritional and diet related outcomes require considerable time in order to produce thorough measurements. Initially the ambition was to collect data on the MAD, minimum dietary diversity for women (MDDW), MDD for children, MMF, Household Diet Diversity Score (HHDDS), Food Consumption Score (FCS) and KAP. After consulting the needs of WFP and hierarchizing the priority indicators related to women's and children's nutrition, HHDDS and FCS have been excluded from the questionnaire.

Another issue pertaining to data collection via phone was the issue of reachability of respondents, both in terms of obtaining the right phone numbers and connectivity. While the former risk was relatively low given that the sampling frame contained updated phone numbers of beneficiaries, we aimed to minimize non-response by engaging woreda officials and health extension workers before the start of the data collection in the verification of phone numbers and reaching out to respondents beforehand to ensure awareness of the upcoming survey.

From an analytical standpoint, descriptive statistics and conclusions must be considered within the COVID-19 pandemic context. Indeed, the sanitary situation potentially alters the capacity of households to access fresh foods during data collection. On the one side, markets might be closed, and attendance might be limited by authorities. Further, we can assume beneficiaries might seek to avoid crowded environments such as the markets. For these reasons, beneficiaries may have used fewer fresh foods and relied more on storable non-perishable foods in order to reduce the frequency of attending markets or leaving the house. Such households would thus consume fewer fresh foods than they would have done without the COVID-19 pandemic.

Tigray Conflict

The ongoing conflict in the Region of Tigray has generated a humanitarian emergency and has had a number of spill-over effects in neighbouring regions such as an influx of thousands of displaced

people and network disruptions.[37] This also raised challenges to the present evaluation data collection and analysis. During the inception phase, the phone network in Seqota woreda was shut down due to the conflict, thereby impeding to conduct phone surveys at that time. After an assessment of the situation, WFP ETHCO and Regional Bureau in Nairobi (RBN) decided on December 2nd that, given the tight timeline for data collection until the end of 2020, this woreda had to be dropped from the sample. Similar to the consequences of the COVID-19 pandemic, the Tigray conflict could have significant effects on access to fresh foods and diet-related practices in the programme areas.

Timeline

Due to WFP funding constraints, data had to be collected within a very short period. The short timeline of data collection available may be associated with reduced quality of collected data mainly because of potential issues identifying and reaching respondents or following up with questionable data. This risk was somewhat reduced through intensive training of enumerators and engaging woreda officials and health extension workers to pre-identify respondents as described above. Another key disadvantage is that the quantitative and qualitative data collection had to be conducted simultaneously, meaning that there was no room to adjust the qualitative tools according to the results from the survey.

Inability to measure impact

Finally, one needs to point out that a proper estimation of the impacts of FFV is not possible. This is due to limited pre-treatment data and the incapacity to select/build a credible counterfactual situation. In particular, we are not able to design an impact evaluation to explore cause and effect questions as there is no adequate information on a credible comparison population currently available.[38,39] We considered a chronological comparison of the outcomes of interest on beneficiaries to illustrate a potential variation throughout time. However, this analysis is dependent on the access to data from the Post Distribution Monitoring surveys conducted since 2019 on a subset of woredas. The current situation does not enable to attribute any variation of the outcomes to WFP's activities. The estimates have to be considered as correlations instead of causal impact measures. On the other hand, an impact evaluation in the current COVID-19 situation would have been of limited usefulness anyhow, in the sense that such evaluation would have been of little guidance for the situation after the pandemic.

Gender dimensions of the sample

GEEW aspects are mainstreamed into different evaluation criteria, particularly relevance and effectiveness. Given that the focus of the programme and the evaluation lies in enhancing nutrition outcomes for women and children, a consensus between the evaluation team and WFP

[37] https://reports.unocha.org/en/country/ethiopia/

[38] WFP was able to retrieve lists of PSNP beneficiaries from neighboring woredas where the programme was not rolled out. However, the key data to determine a credible comparable control group based on eligibility criteria as of year 2018 or 2019 are missing. The database contains information as of today, but not about the population that would have been eligible in 2018 or 2019.

[39] We also explored a possible before-after comparison design for key outcomes of beneficiaries in order to shed light on the performance of beneficiaries with respect to nutrition outcomes targeted by the implementation. WFP has been conducting annual PDM surveys with small samples of beneficiaries since 2019 in a subset of woredas, including information on food-security and nutrition of mothers and their children. However, the compiled PDM databases were not complete and therefore not usable for sampling purposes.

was reached that women will be the sole interview respondents. While this allows for a rich sample of nutrition outcomes for women and their children (boys and girls), the analysis of attitudes and practices related to household food consumption and child feeding is limited to the female perspective only. The lack of information from male household members does not allow to investigate household decision-making.

5. Survey Implementation

The quantitative data collection was implemented by a survey team consisting of one field coordinator, two supervisors and 14 enumerators. The training of supervisors and enumerators took place between 10 and 14 December 2020. On the first day, the supervisors were trained on the project, survey instruments, roles and responsibilities, and phone survey protocols. A three-day enumerator training was then conducted with 15 enumerator candidates to familiarize them with the tools and processes of the phone survey. The training was conducted remotely by C4ED with organizational support by the local partner. The training participants gathered in a hotel room under Covid-19 safety measures (e.g. wearing of masks at all times and frequent ventilation of the room). On the fourth training day, a pilot was conducted with households from the beneficiary list outside of the sample. Each enumerator conducted two pilot interviews. The pilot was followed by a debrief session where issues and experiences during the interviews were shared and programming errors rectified. Based on the performance throughout the training and the pilot, 14 enumerators and one backup enumerator were selected. The survey team was comprised of male and female enumerators.

The surveys with households and traders were conducted between 15 December 2020 and 08 January 2021. The teams were given lists of names, phone numbers and location (kebele or village) of the respondents. One week before the start of the data collection, supervisors travelled to the woredas to coordinate with woreda officers and health extension workers a strategy to reach the beneficiaries. Beforehand, a concern was raised that beneficiaries may turn on their phones only infrequently or during market days. Therefore, health extension workers identified beneficiaries in their kebeles and made appointments for them to gather at a central location or instructed them to turn on their phones. This process was ongoing during the data collection. The efforts of health extension workers were incentivized with a small Birr value. This engagement undoubtedly led to high response rates of phone interviews.

The data collection took place remotely with enumerators and supervisors located together in Addis Ababa. Interview languages included Amharic and Oromoffia.

During data collection, the incoming data quality was constantly assessed through high-frequency checks by the C4ED team, including inconsistencies, outliers, and other potential entry errors. We also checked completeness of submitted forms, duplicates in identifying variables, and missingness of key variables. This allowed for real-time feedback and rectification of errors. If inconsistencies were detected, the enumerators were asked to clarify the situation including call-back of respondents. Most issues could be credibly resolved, while only in some cases the decision was made to drop and replace the interview.

For the IDI, three Ethiopian qualitative researchers were recruited to conduct the 28 interviews. A one-day training on 29 December 2020 at Addis Ababa University familiarized the experienced researchers with the tools and the remote data collection processes. The interviews were conducted between 31 December 2020 and 19 January 2021.

Annex 7: Ethical Considerations

Table A13: Ethical Considerations

Stage	Ethical concern	Safeguarding Measure
Preparation/design	Cultural sensitivity of questions	Review of tools by WFP and Evaluation Reference Group members An international Institutional Review Board approval has been sought and data collection will only start once ethical approval has been granted Questionnaire translated into local languages (Amharic and Oromiffaa) and pre-tested twice (during supervisor and enumerator training)
Data collection	Health risks related to COVID-19	Both quantitative and qualitative data will be collected via phone, such that no physical contact is required between enumerators and respondents
	Proper enumerator conduct (e.g. ensuring informed consent, protecting privacy, confidentiality and anonymity of participants)	Employ experienced enumerators Intensive training of enumerators In-class and in-field practice sessions Close supervision of conduct by supervisor and field coordinator
Data analysis, reporting and dissemination	Privacy, confidentiality and anonymity of participants	Data will be stored on C4ED's SurveyCTO server which can only be accessed by the evaluation team. Before analysis and data sharing, the data will be anonymized, i.e. all personally identifiable information such as name and phone number will be removed.

Annex 8: Descriptive statistics

Table A 14: Descriptive statistics for households with low and high FFV redemptions

Variable	(1) Less than 4 redemptions N	Mean/SE	(2) At least 4 redemptions N	Mean/SE	T-test Difference (1)-(2)
Dummy for child in HH <6 months	350	0.071 (0.014)	1019	0.069 (0.008)	0.003
Dummy for child in HH 6-23 months	350	0.720 (0.024)	1019	0.639 (0.015)	0.081***
Dummy for child in HH 24-48 months	350	0.383 (0.026)	1019	0.499 (0.016)	-0.116***
Dummy for child in HH >48 months	350	0.003 (0.003)	1019	0.002 (0.001)	0.001
Season of registration	351	0.932 (0.013)	1029	0.747 (0.014)	0.184***
Woreda: Dawa-Cheffa	351	0.168 (0.020)	1029	0.166 (0.012)	0.002
Woreda: Dessie-Zuria	351	0.057 (0.012)	1029	0.207 (0.013)	-0.150***
Woreda: Habru	351	0.140 (0.019)	1029	0.174 (0.012)	-0.034
Woreda: Kalu	351	0.259 (0.023)	1029	0.135 (0.011)	0.124***
Woreda: Kobo	351	0.157 (0.019)	1029	0.169 (0.012)	-0.012
Woreda: Mekdella	351	0.219 (0.022)	1029	0.149 (0.011)	0.071***
Registered in 2019	351	0.647 (0.026)	1029	0.450 (0.016)	0.197***
Pregnant	351	0.313 (0.025)	1029	0.563 (0.015)	-0.249***
Married	351	0.932 (0.013)	1029	0.950 (0.007)	-0.019
Mother has formal education	351	0.487 (0.027)	1029	0.550 (0.016)	-0.063**
Distance to nearest market (in minutes)	345	77.223 (3.126)	1008	65.169 (1.712)	12.055***
Access to Markets during COVID-19	342	0.661 (0.026)	1009	0.653 (0.015)	0.008
Number of HH members	351	5.536 (0.097)	1029	5.659 (0.057)	-0.123
Number of children in HH	351	1.182 (0.021)	1028	1.239 (0.015)	-0.057**
HH has electricity	351	0.248 (0.023)	1029	0.300 (0.014)	-0.052*
HH owns milk cows	351	0.456 (0.027)	1029	0.493 (0.016)	-0.037
HH owns oxen or bulls	351	0.501 (0.027)	1029	0.497 (0.016)	0.005
HH owns Chicken	351	0.516 (0.027)	1029	0.559 (0.015)	-0.043

Annex 9: Regression approach and results

Child nutrition outcomes

To further examine mechanisms that influence each respective nutrition outcome, we ran several multivariate regressions, including a number of possible explanatory factors at the level of children, household and market. The estimations were carried out for children aged 6 to 23 months in households registered during 2018 or 2019 and still actively enrolled in the programme (as of December 2020). We estimated for each nutrition outcome the following estimation model:

$$Y_{ijk} = \beta_0 + \beta_1 Age_i + \beta_2 {Age_i}^2 + \beta_3 {Age_i}^3 + \beta_4 BirthSeason_i + \beta_5 Female_i + \beta_6 Breastfed_i + \gamma Woreda_k + \delta {X_j}^{Household} + \theta {X_j}^{Program} + \epsilon_i \quad \text{(Eq. 1)}$$

Y_{ijk} refers to one of the three nutrition outcomes, i.e. the MAD, MDD and MFF for child i observed in household j living in woreda k. The first characteristic we included was the age of the child, calculated in months. We also introduced the quadratic and cubic term of age, in order to take into account possible non-linear relationships between these characteristic and nutrition outcomes. $BirthSeason_i$ refers to the season of birth of the child, distinguishing between being born during the harvest season (October to May) or the lean season (June to September). $Female_i$ is an indicator of the child being female, with the alternative being male, and $Breastfed_i$ is an indicator of the child being breastfed at the time of the interview. $Woreda_k$ is an indicator of the household living in Woreda k, where the omitted Woreda is *Dawa-Cheffa*. Note that we do not include an indicator for the year of registration of the household. This is because all households in a given woreda were registered in the same year (hence such indicator would be collinear with the woreda indicator)[40]. The regression is weighted by sample probability weights based on the sampling strategy adopted during the end-line data collection. This ensured that the estimated coefficients are representative of the overall beneficiary population. Standard errors are robust to heteroscedasticity and clustered at the household level to account for the correlation of outcomes within the same households.

Further, we considered a set of household characteristics ${X_j}^{Household}$. This set comprises education, pregnancy and marriage status of the mother, number of household members and number of children living in the household, as well as proxy indicators of the household's wealth, i.e. having access to electricity, owning different common kinds of livestock. Also included in ${X_j}^{Household}$ are indicators for a household's distance to the nearest market, measured in minutes of travel with the household's usual means of transport, as well as an indicator for the household having unrestricted access to markets during the COVID-19 pandemic.

Lastly, we include programme-specific variables in ${X_j}^{Program}$, consisting of an indicator for households having received transfers from the PSNP during the last 12 months, as well as an indicator for the households having received a fresh food voucher during the last 6 months.

[40] Households in Woredas Dessie-Zuria, Habru and Kobo were registered during 2018; households in Woredas Dawa-Cheffa, Kalu and Mekdella were registered during 2019.

Estimates of the regression model are reported in Table A 15.**Error! Reference source not found.**[41]

Table A 15: Main Regression Estimates for Nutrition Outcomes of Children (6-23 months)

	MAD			MMF			MDD		
	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 1	(5) Model 2	(6) Model 3	(7) Model 1	(8) Model 2	(9) Model 3
Age in months	0.405**	0.393*	0.381*	-0.049	-0.020	-0.019	0.414**	0.420**	0.409**
	(0.156)	(0.153)	(0.153)	(0.073)	(0.075)	(0.074)	(0.148)	(0.145)	(0.145)
Age in months (squared)	-0.029**	-0.028*	-0.027*	0.004	0.002	0.002	-0.030**	-0.030**	-0.029**
	(0.011)	(0.011)	(0.011)	(0.005)	(0.005)	(0.005)	(0.011)	(0.011)	(0.011)
Age in months (cubic)	0.001**	0.001*	0.001*	-0.000	-0.000	-0.000	0.001**	0.001**	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Birth in lean season	0.034	0.082	0.083	-0.086*	-0.079	-0.076	0.052	0.104	0.106
	(0.068)	(0.065)	(0.064)	(0.036)	(0.041)	(0.041)	(0.066)	(0.061)	(0.060)
Woreda: Dessie-Zuria	0.348***	0.178	0.162	-0.092*	-0.157*	-0.147*	0.357***	0.293*	0.276*
	(0.089)	(0.126)	(0.123)	(0.037)	(0.065)	(0.067)	(0.088)	(0.130)	(0.125)
Woreda: Habru	0.342**	0.216	0.224	-0.136*	-0.171*	-0.167*	0.386***	0.365**	0.374**
	(0.118)	(0.133)	(0.134)	(0.065)	(0.077)	(0.078)	(0.112)	(0.133)	(0.133)
Woreda: Kalu	0.277***	0.194*	0.179*	-0.127**	-0.152***	-0.153***	0.331***	0.260**	0.242**
	(0.073)	(0.080)	(0.081)	(0.039)	(0.046)	(0.044)	(0.072)	(0.081)	(0.082)
Woreda: Kobo	0.281**	0.144	0.144	-0.095*	-0.150*	-0.146*	0.361***	0.324**	0.324**
	(0.086)	(0.106)	(0.106)	(0.042)	(0.061)	(0.061)	(0.085)	(0.110)	(0.109)
Woreda: Mekdella	0.134	0.014	-0.001	-0.038	-0.085*	-0.085*	0.129	0.056	0.038
	(0.070)	(0.081)	(0.082)	(0.020)	(0.036)	(0.035)	(0.075)	(0.085)	(0.086)
Female	-0.038	-0.031	-0.030	-0.004	-0.008	-0.010	0.015	0.015	0.017
	(0.051)	(0.049)	(0.049)	(0.028)	(0.029)	(0.029)	(0.051)	(0.048)	(0.048)
Child is currently breastfed	0.259**	0.130	0.131	0.009	0.012	0.015	-0.091	-0.168	-0.169
	(0.089)	(0.085)	(0.085)	(0.028)	(0.031)	(0.031)	(0.098)	(0.087)	(0.087)
Pregnant		0.060	0.053		0.010	0.007		-0.026	-0.035
		(0.077)	(0.076)		(0.039)	(0.040)		(0.077)	(0.076)
Married		0.205**	0.214**		-0.113**	-0.117**		0.327***	0.337***
		(0.073)	(0.072)		(0.037)	(0.038)		(0.075)	(0.075)
Mother has formal education		-0.045	-0.041		-0.020	-0.023		-0.025	-0.021
		(0.052)	(0.052)		(0.027)	(0.027)		(0.050)	(0.050)
Number of HH members		-0.009	-0.010		0.003	0.002		-0.013	-0.014
		(0.016)	(0.016)		(0.008)	(0.008)		(0.016)	(0.016)
Number of children in HH		-0.120*	-0.124*		-0.024	-0.023		-0.107	-0.111
		(0.061)	(0.060)		(0.032)	(0.031)		(0.063)	(0.063)
HH has electricity		-0.012	-0.009		-0.055	-0.055		-0.007	-0.004
		(0.063)	(0.063)		(0.038)	(0.039)		(0.060)	(0.060)
HH owns milk cows		0.105	0.111*		0.086*	0.086*		0.111*	0.118*
		(0.056)	(0.056)		(0.035)	(0.035)		(0.053)	(0.053)
HH owns oxen or bulls		0.100	0.104		-0.009	-0.009		0.051	0.057
		(0.060)	(0.060)		(0.039)	(0.039)		(0.058)	(0.057)
HH owns Chicken		0.041	0.035		0.017	0.017		0.041	0.034
		(0.052)	(0.052)		(0.029)	(0.029)		(0.050)	(0.050)
Market distance: 30 - 60 minutes		-0.102	-0.100		0.040	0.043		-0.103	-0.100
		(0.066)	(0.067)		(0.036)	(0.036)		(0.066)	(0.066)
Market distance: 1 - 2 hours		-0.090	-0.090		-0.002	0.000		-0.086	-0.086
		(0.072)	(0.072)		(0.034)	(0.034)		(0.066)	(0.066)
Market distance: Over 2 hours		-0.054	-0.067		-0.036	-0.036		-0.098	-0.114
		(0.084)	(0.085)		(0.037)	(0.037)		(0.081)	(0.081)
Access to Markets during COVID-19		0.015	0.006		0.034	0.033		0.012	0.000
		(0.054)	(0.054)		(0.030)	(0.031)		(0.052)	(0.052)
Transfer from PSNP			0.071			-0.005			0.086
			(0.062)			(0.040)			(0.059)
FFV received in last 6 months			0.108			0.098			0.163
			(0.207)			(0.108)			(0.214)
Constant	-1.748**	-1.603*	-1.658*	1.133***	1.110**	1.015**	-1.517*	-1.526*	-1.633*
	(0.655)	(0.657)	(0.697)	(0.341)	(0.369)	(0.388)	(0.637)	(0.643)	(0.687)
Observations	692	678	678	702	688	688	702	688	688

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Considering the full model of the MAD in column (3) we find that the average MAD score significantly increases with the age of the child. Furthermore, the quadratic and the cubic term of

[41] For each outcome (MAD, MFF and MDD), we estimate three different models, with a stepwise increase in included explanatory variables. The first model for each nutrition outcome, that is columns (1), (4) and (7) present estimates of the basic model specification, including only child-specific and woreda variables. The second model for each nutrition outcome, that is columns (2), (5) and (8) present estimates of an extended model including households' characteristics, as described above. The full model specification is presented in the third model for each nutrition outcome, that is columns (3), (6) and (9), including program-specific variables in addition to child- and woreda specific characteristics and households characteristics.

age are also significant. The first indicates a concave relationship between the nutrition outcome and age of the child, i.e. diminishing effect of age on the MAD score as age increases. The positive estimate of the cubic term indicates how this diminishing effect slows down in the higher regions of age (up to 24 months). Considering age coefficients for MDD and MFF, it appears that the impact of age on the MAD is mainly driven by the MDD.

We further find that while woreda-specific indicators are highly significant in the basic model in column (1) of the MAD, while this effect diminishes when including household- and program-specific covariates in models 2 and 3 of the MAD. Among household characteristics, the mother being married is significantly associated with higher likelihood of the child meeting MAD requirements. This effect differs in sign between the MMF, where being married has a negative effect on the outcome, and the MDD, where being married has a positive effect on the outcome.

Children that were breastfed at the time of interview are more likely to meet MAD requirements, but this effect is not statistically significant for the MDD and MFF outcomes. Children in households owning milk cows have a significantly larger likelihood of meeting requirements of the MAD, MDD and MFF. The likelihood of meeting MAD requirements of a child decreases with the number of children in the household. This effect is not statistically significant for the MDD and MFF outcomes. We do not find any significant associations of PSNP transfers or reception of fresh food vouchers with the MAD indicator.

Nutrition scores for women

To further examine possible mechanisms and characteristics by which the MDD-W outcome for women may be affected, we implemented a multivariate regression including a number of explanatory factors at the level of household and market, similar to the regression model for children's nutrition**Error! Reference source not found.**. The estimations are carried out for w omen aged 18 to 49 years who registered in 2018 or 2019 and who have been enrolled in the program as of December 2020. We estimate for the MDD-W outcome the following estimation model:

$$MDDW_{jk} = \beta_0 + \beta_1 YoungestChild_j + \beta_2 RegSeason_j + \beta_2 Breastfeeding_j + \beta_3 Pregnant_j + \beta_4 MDDWEvent_j + \gamma Woreda_k + \delta X_j^{Household} + \theta X_j^{Program} + \epsilon_i \quad \text{(Eq. 2)}$$

$MDDW_{jk}$ refers to the Minimum Diet Diversity outcome for the observed woman in household j living in woreda k. The first characteristic we include is an indicator for the age group of the youngest child living in the household, with age groups being less than 6 months, 6-23 months and 24-48 months. The omitted age group for the estimation are children less than 6 months old. $RegSeason_j$ refers to the season in which household j was registered, $Breastfeeding_j$ is an indicator showing whether the observed woman in household j is currently breastfeeding and $Pregnant_j$ is an indicator of the mother in household j being pregnant. $MDDWNormal_j$ is an indicator for any event or problem the day before the interview that impacted the beneficiaries normal eating habits. $Woreda_k$ is an indicator of the household living in Woreda k, where the omitted Woreda is *Dawa-Cheffa*. The regression is weighted by sample probability weights based on the sampling strategy adopted during the end-line data collection. Further, we consider a set

of household characteristics $X_j^{Household}$ and a set of programme-specific variables in $X_j^{Program}$. Both sets of variables are identically defined as beforein the estimation model for children).

The next table reports estimate of the regression model for the MDD-W.[42] We find that women whose youngest child is between 24 and 48 months old are significantly less likely to meet MDD-W requirements compared to women whose youngest child is less than 6 months old. Furthermore, the total number of children in a household significantly decreases a woman's likelihood to meet MDDW requirements. Women in woreda Kobo are significantly less likely to meet MDD-W requirements. This effect persists even when accounting for a number of household- and program-specific variables. Further, we find that women in households that own oxen or bulls are significantly more likely to meet MDD-W requirements. The likelihood of a woman meeting MDD-W requirements decreases significantly if the household faced restricted access to markets during the COVID-19 pandemic. Having received an FFV in the last 6 months significantly increases the likelihood of women meeting MDD-W requirements.

Table A 16: Main Regression Estimates for Minimum Diet Diversity of Women aged 15-49

	MDDW		
	(1) Model 1	(2) Model 2	(3) Model 3
Youngest child: 6-23 months	-0.095	-0.170	-0.177
	(0.095)	(0.095)	(0.094)
Youngest child: 24-48 months	-0.179	-0.283*	-0.292*
	(0.126)	(0.132)	(0.130)
Registered in harvest season	-0.002	-0.015	-0.006
	(0.074)	(0.072)	(0.073)
Woreda: Dessie-Zuria	-0.088	-0.148	-0.154
	(0.085)	(0.090)	(0.093)
Woreda: Habru	-0.116	-0.157	-0.136
	(0.097)	(0.095)	(0.097)
Woreda: Kalu	0.180*	0.143	0.128
	(0.082)	(0.080)	(0.081)
Woreda: Kobo	-0.174*	-0.269***	-0.273***
	(0.080)	(0.077)	(0.077)
Woreda: Mekdella	0.036	0.022	0.005
	(0.079)	(0.080)	(0.082)
Child is currently breastfed	0.003	-0.056	-0.060
	(0.063)	(0.063)	(0.063)
Pregnant	-0.011	0.013	0.007
	(0.054)	(0.051)	(0.051)
Some issue affected yesterday's eating		-0.099	-0.092
		(0.058)	(0.058)
Married		0.013	0.012
		(0.081)	(0.085)
Mother has formal education		-0.007	-0.006
		(0.047)	(0.048)
Number of HH members		0.012	0.010
		(0.013)	(0.013)
Number of children in HH		-0.145**	-0.146**
		(0.051)	(0.052)
HH has electricity		0.057	0.059
		(0.054)	(0.054)
HH owns milk cows		-0.002	0.007
		(0.050)	(0.050)
HH owns oxen or bulls		0.143**	0.147**
		(0.053)	(0.053)
HH owns Chicken		0.076	0.070
		(0.048)	(0.048)

[42] We estimate three different models, with a stepwise increase in included explanatory variables. Column (1) presents estimates of the basic model specification, including beneficiary-specific and woreda variables. Column (2) presents estimates of an extended model including household characteristics. Column (3) presents estimates of the full model including program-specific variables in addition to child- and woreda specific characteristics and household characteristics.

Market distance: 30 - 60 minutes		-0.078	-0.078
		(0.060)	(0.060)
Market distance: 1 - 2 hours		-0.107	-0.109
		(0.065)	(0.065)
Market distance: Over 2 hours		-0.118	-0.134
		(0.070)	(0.069)
Access to Markets during COVID-19		-0.093	-0.101*
		(0.050)	(0.050)
Transfer from PSNP			0.066
			(0.058)
FFV received in last 6			0.164*
			(0.075)
months Constant	0.417***	0.676***	0.528**
	(0.122)	(0.191)	(0.194)
Observations	697	687	686

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Annex 10: Age groups in graduated Households

Table A 17: Age groups in graduated households

Age groups in graduated HHs	Number of children	% of children
below 6 months	45	6.49
6 to 23 months	215	31.02
24 to 48 months	433	62.48
Total	**693**	**100**

Annex 11: Regression results for graduated households

Table A 18 shows results of the regression estimation, where columns (1) and (2) show a stepwise increase in covariates for enrolled children aged 6-23 months, and columns (3) and (4) show a stepwise increase in covariates for graduated children aged 24-48 months.

Table A 18: Regression estimates for enrolled children (6-23 months) and graduated children (24-48 months)

	Enrolled (6-23 months)		Graduated (24-48 months)	
	(1)	(2)	(3)	(4)
	Model 1	Model 2	Model 1	Model 2
Age in months	0.414**	0.409**	1.023*	0.996**
	(0.148)	(0.145)	(0.428)	(0.380)
Age in months (squared)	-0.030**	-0.029**	-0.030*	-0.029**
	(0.011)	(0.011)	(0.013)	(0.011)
Age in months (cubic)	0.001**	0.001**	0.000*	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Birth in lean season	0.052	0.106	-0.062	-0.145*
	(0.066)	(0.060)	(0.072)	(0.064)
Woreda: Dessie-Zuria	0.357***	0.276*	0.180	-0.021
	(0.088)	(0.125)	(0.152)	(0.144)
Woreda: Habru	0.386***	0.374**	0.143	-0.007
	(0.112)	(0.133)	(0.154)	(0.150)
Woreda: Kalu	0.331***	0.242**	-0.004	-0.150
	(0.072)	(0.082)	(0.179)	(0.178)
Woreda: Kobo	0.361***	0.324**	0.092	-0.104
	(0.085)	(0.109)	(0.152)	(0.149)
Woreda: Mekdella	0.129	0.038	0.019	-0.059
	(0.075)	(0.086)	(0.161)	(0.155)
Female	0.015	0.017	0.044	0.066
	(0.051)	(0.048)	(0.063)	(0.058)
Child is currently breastfed	-0.091	-0.169	0.184**	0.159*
	(0.098)	(0.087)	(0.071)	(0.075)
Pregnant		-0.035		0.161
		(0.076)		(0.087)
Married		0.337***		0.273**
		(0.075)		(0.102)
Mother has formal education		-0.021		0.223***
		(0.050)		(0.062)
Number of HH members		-0.014		-0.007
		(0.016)		(0.021)
Number of children in HH		-0.111		0.050
		(0.063)		(0.079)
HH has electricity		-0.004		0.072
		(0.060)		(0.075)
HH owns milk cows		0.118*		0.133
		(0.053)		(0.068)
HH owns oxen or bulls		0.057		-0.032
		(0.057)		(0.072)
HH owns Chicken		0.034		0.275***
		(0.050)		(0.063)
Market distance: 30 - 60 minutes		-0.100		-0.076
		(0.066)		(0.073)
Market distance: 1 - 2 hours		-0.086		0.012
		(0.066)		(0.084)
Market distance: Over 2 hours		-0.114		-0.090
		(0.081)		(0.102)
Access to Markets during COVID-19		0.000		0.167*
		(0.052)		(0.068)
Transfer from PSNP		0.086		-0.140
		(0.059)		(0.078)
FFV received in last 6 months		0.163		0.082
		(0.214)		(0.072)
Constant	-1.517*	-1.633*	-10.928*	-11.292**
	(0.637)	(0.687)	(4.763)	(4.266)
Observations	702	688	413	382

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Annex 12: Bibliography

Bose I., K. L. (2019). The "Fill the Nutrient Gap" analysis: An approach to strengthen nutrition situation analysis and decision-making towards multi-sectoral policies and systems change. *Matern Child Nutr*.

CSA and ICF. (2016). *Ethiopia Demographic and Health Survey 2016.* Addis Ababa, Ethiopia, and Rockville, Maryland, USA: Central Statistical Agency (CSA) Ethiopca and ICF.

EPHI and WFP. (2013, June). The Cost of Hunger in Ethiopia: Implications for the Growth and Transformation of Ethiopia.

EPHI and WFP. (2020, December). Fill the Nutrient Gap Ethiopia. Summary of Results Report (Working Draft V2.0).

FAO, IFAD, UNICEF, WFP and WHO. (2020). *The State of Food Security and Nutrition in the World 2020. Transforming food systems for affordable healthy diets.* Rome: FAO.

Federal Republic of Ethiopia. (2016). National Nutrition Program 2016-2020.

Federal Republic of Ethiopia. (2016, March). Seqota Declaration Implementation Plan (2016-2020): Summary Programme Approach Document.

Federal Republic of Ethiopia. (2018). Food and Nutrition Policy.

Hirvonen K, a. W. (2019). Consumption, production, market access and affordability of nutritious foods in the Amhara Region of Ethiopia. Addis Ababa, Ethiopia: Alive & Thrive and Interntaional Food Policy Research Institute .

Mekonnen, D. T. (2020). Can household dietary diversity inform about nutrient adequacy? Lessons from a food systems analysis in Ethiopia. . *Food Sec*, 1367-1383.

MOWCY, UNICEF Ethiopia and SPRI. (2019). Gender Equality, Women's Empowerment and Child Wellbeing in Ethiopia.

OCHA. (2020, November 14). Ethiopia: Tigray Region Humanitarian Update. Situation Report No. 5. Ethiopia: United Nations Office for the Coordination of Humanitarian Affairs.

OPHI. (2020, May 11). *Global MPI Country Briefing 2019: Ethiopia (Sub-Saharan Africa)*. Retrieved from OPHI: https://ophi.org.uk/wp-content/uploads/CB_ETH_2019_2.pdf accessed 11/5/2020

P., S. (n.d.). *Food rations cut for refugees in Eastern Africa as coronavirus stretches resources*. Retrieved from WFP: https://www.wfp.org/news/food-rations-cut-refugees-eastern-africa-coronavirus-stretches-resources

UNICEF . (2020, August). Ethiopia Country Office - Humanitarian Situation Report (April to May 2020). Ethiopia.

UNICEF. (2019). *Situation Analysis of Children and Women: Amhara Region.* UNICEF Ethiopia.

WFP. (2020, November 18). *Ethiopia*. Retrieved from https://www.wfp.org/countries/ethiopia

List of Acronyms

BMZ	German Federal Ministry for Economic Cooperation (Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung)
C4ED	Center for Evaluation and Development
DA	Development Agents
DAC	Development Assistance Committee
EQ	Evaluation Question
ETHCO	Ethiopia Country Office
EPHI	Ethiopian Public Health Institute
FCS	Food Consumption Score
FFV	The Fresh Food Voucher
FNG	Fill the Nutrient Gap
FNP	Food and Nutrition Policy
GEEW	Gender Equality and the Empowerment of women
GOE	Government of Ethiopia
HDA	Health Development Army
HEW	Health Extension Workers
HHDDS	Household Diet Diversity Score
IDI	In-Depth Interviews
IYCF	Infant and Young Child Feeding
KAP	Knowledge, Attitudes and Practices
KFW	German Development Bank (Kreditanstalt für Wiederaufbau)
KII	Key Informant Interviews
MAD	Minimum Acceptable Diet
MDD	Minimum Diet Diversity
MDD-W	Minimum Diet Diversity for Women
MMF	Minimum Meal Frequency
NNP	National Nutrition Programme
NNS	National Nutrition Strategy
PDM	Post-Distribution Monitoring
PLW	Pregnant and lactating women
PSNP	Production Safety Net Programme
QA	Quality Assurance
RBN	Regional Bureau Nairobi
SBCC	Social and Behaviour Change Communication
USAID	United States Agency for International Development
WDA	Women's Development Army
WFP	World Food Programme
WHO	World Health Organisation

[Place, Month and Year, Report number]

[Name of commissioning Office]
[Link to the website]