Freie Universität Berlin

Department of Economics

The Impact of Small-Scale Agricultural Support on Household Food Security in Least Developed Settings Evidence from South Sudan[1]

A thesis presented for the degree of Master of Science in Public Economics.

Dorothee Weiffen

September 2021

Supervisors:

Prof. Dr. Natalia Danzer[2]

Prof. Dr. Tilman Brück[3]

http://dx.doi.org/10.17169/refubium-36207

[1]**Acknowledgements:** I thank Dr. Ghassan Baliki for the brilliant inspiration and tireless support in my academic development, Prof. Natalia Danzer, and Prof. Tilman Brück for valuable input and feedback, IGZ and ISDC for the opportunity to collaborate in this project. I thank Jean-Paul Kadigi, Evance Lisok, Alfred Izama, and Anja Müting from Malteser for their collaboration, effort, and reachability. I thank the enumerators for their valuable work and the study participants for their time and confidence.

[2] Department of Economics, Free University Berlin, Germany

[3] Group for Economic Development and Food Security, Leibniz-Institute for Vegetable and Ornamental Crops

Structure

Abbreviations

ACLED	Armed Conflict Location and Event Data Project
ATT	Average Treatment Effect on the Treated
CG	Control Group
DiD	Difference-in-Differences
DiDiD	Difference-in-Differences-in-Differences
FAO	Food and Agriculture Organization of the United Nations
FCS	Food Consumption Score
FHI	Family Health International
FIES	Food Insecurity Experience Score
FSIN	Food Security Information Network
HDX	Humanitarian Data Exchange
HFIAS	Household Food Insecurity Access Scale
HH	Household
HHH	Household Head
IDP	Internally Displaced Person
IGZ	Leibniz Institute of Vegetable and Ornamental Crops
IMF	International Monetary Fund
IPC	Integrated Food Security Phase Classification
MDD	Minimum Dietary Diversity Score
MI	Malteser International
NGO	Non-governmental Organization
OCHA	Office for the Coordination of Humanitarian Affairs of the United Nations
OLS	Ordinary Least Squares
PSM	Propensity Score Matching
RCT	Randomized Control Trial
SDG	Sustainable Development Goals
SE	Standard Errors
SAI	Small-Scale Agricultural Intervention
TG	Treatment Group
WASH	Water, Hygiene, and Sanitation
WFP	World Food Programme

Tables

Figures

Non-technical Summary

The second Sustainable Development Goal of the United Nations (SDG 2) sets the goal of achieving zero hunger worldwide by 2030. Still in 2019, around 750 million people did not have secure access to adequate food (FAO et al., 2020). People living in underdeveloped regions affected by conflicts are especially threatened by limited access to an adequate food quality and quantity. Smallholding farmers in underdeveloped areas tend to be extremely poor and are particularly at risk of unexpected events. Events like theft, floods, droughts, high prices of seeds, livestock or crop pests, unforeseen high expenditures, like a disease of a household member reduce their access to food from homestead production or induce financial losses because these people often have neither insurances nor savings. The permanent risk of these challenges combined to poverty and conflict exacerbate the sparing nutrition of farmers in underdeveloped regions.

To improve these people's access to a better nutrition and to realize SDG 2, many organizations aid smallholder farmers. Support frequently includes the provision of agricultural training, seeds, or tools to enhance the agricultural production and increase like this the people's food access. Alternatively, money or food can be distributed to mitigate hunger directly. To justify these aid programs to their donors and to spend the funds in the best manner possible, organizations need proof of the success of such programs. However, little is known about the effectiveness in challenging settings affected by conflicts and unforeseen events. Existing research highlights a strong potential of such programs to increase harvests and nutrition of farmers in peaceful settings; however, it remains to be investigated if these are still effective under extreme conditions like conflicts and severe agricultural challenges. Further, it is vital to know whether agricultural support methods or direct food or money provision programs are more effective in enhancing people's food security.

This thesis concentrates on these questions. It examines the effectiveness of such a support program in South Sudan after one year of implementation with modern statistical methods. In South Sudan, 53% of the people were insufficiently nourished in 2020 (IPC Global Partners, n.d.). Additionally, the people suffered through a civil war from 2013 to 2018 and are still exposed to violence to a high degree. A frequent occurrence of severe events complicates the peoples' nutritional situation further: In 2020, a severe flood displaced 500 thousand people in the country (OCHA, 2021), and agricultural activities are additionally challenged by livestock diseases and locust infestations. The high degree of migration and displacement fosters land conflicts (McMichael, 2016). In 2020, nearly 2 million people have been internally displaced (World Bank, 2020) and 2.2 million have fled to neighboring countries (UNHCR, 2021). Furthermore, the economy of the country was widely impacted by the Covid-19 pandemic and an unstable political situation.

The aid program that is evaluated in this work started in 2019 and runs for three years. The households received agricultural training sessions, seeds, and tools. Furthermore, participating households that were affected by the flood catastrophe received a one-time cash transfer. The goal is to enhance and diversify homestead food production, not only to ensure the households' constant access to adequate food but also to make them less dependent on market prices and to buffer challenges like climatic events.

For this study, 847 households were surveyed twice: before and during the aid program. Around half of the respondents received the support. Within the surveyed households, a high degree of exposure to severe challenges and a high rate of food insecure households were detected. The analysis reveals that after one year, respondents, who participated in the aid program, increased their dietary diversity compared to non-participating respondents. On average, the participating respondents ate one additional type of food per day because of the program. The respondents adjust their food choices away from green leafy vegetables towards other vegetables, nuts, pulses, meat, poultry, fish, and eggs. Another measure that captures the households' dietary diversity instead of the respondents' dietary diversity also indicates changes in food choices, however, no influence on the overall number of consumed food groups. The last main measure captures if the respondents were exposed to situations, in which adequate nutrition was limited. The analysis does not stress changes in the total number of experienced situations due to the program. However, compared to respondents, who did not participate in the support program, the program helped the participating respondents to decrease the probability to have passed a whole day without food within one year.

Within one year, the participants' harvest of cereals, other main crops like pulses and nuts, vegetables, and fruits remained roughly at a constant level while the harvest of non-participating households decreased drastically, so that the agricultural program buffered the challenging environmental influences that would have decreased the participants' harvest. The cash transfer to beneficiary households, who were severely affected by the flood during the study year combined to the agricultural support, induced the improvements in dietary diversity. The agricultural program alone could not improve the diet diversity without the cash transfer. This also means that the transfer probably balanced out the negative consequences of the flood.

Being severely exposed to challenges that affect a larger fraction of the population did not hinder the effect of the aid program on nutrition because the whole population on the study-site is frequently exposed to such events. If households experience a severe illness or an accident of the household head, the positive effect of the program on dietary diversity disappears.

After one year of program implementation, which was shaded by a range of challenging obstacles, the program shows a buffering effect on their agricultural production. Therefore, existing knowledge of such interventions in peaceful settings is comparable to the present case in a conflict-affected region regarding agricultural productivity. Concerning food security, substantial improvements like most of the other literature in peaceful settings, is not assessed after one year. A larger impact in the following two years of the program is expected. Especially, if the conflict- and the Covid-19-situation ease, the program is expected to yield further progress for the participating households.

1 Introduction

SDG 2 sets the goal of achieving zero hunger worldwide by the year 2030. Progress has been made in the past two decades. From 2005 to 2019, the share of undernourished people decreased from 12.6% in 2005 to 8.9% worldwide (FAO et al., 2020). Still, around 750 million people faced severe food insecurity in 2019 (ibid.). Out of the ten worst-affected countries by food crises, six are experiencing severe protracted conflict (FSIN, 2021). Further, in countries that are prone to climatic shocks and locust plagues, people are particularly threatened by famine, mal-, and undernutrition (United Nations, 2020a). Farmers in least developed countries are especially vulnerable to these harvest-affecting events. Therefore, farming households in conflict-affected regions with a high risk of adverse shocks are highly exposed to food insecurity.

South Sudan is one of the least developed countries in the world (United Nations, 2020b). 53% of the population is affected by acute food insecurity in 2020 (IPC Global Partners, n.d.). In combination to severe food insecurity, households are permanently exposed to a variety of threats. Most pressingly are the persistent high levels of violence and the recurrent conflicts (Amnesty International, 2020), the extreme weather conditions (Quinn et al., 2019), and the unstable political and economic situation (World Bank, 2021a). More than 2.2 million South Sudanese people find themselves displaced thus far in 2021 (UNHCR, 2021). In 2021, 82% of the population face absolute poverty (World Bank, 2021a).

One pathway to address food security is through enhancing the productivity of poor farmers. Small-scale agricultural interventions (SAI) are one of many development agendas to contribute to achieving SDG 2 (Abraham and Pingali, 2020). The programs usually strengthen the access to agricultural inputs and provide knowledge transfers to enhance production. The literature provides evidence that diverse forms of SAI can positively influence nutrition in more peaceful developmental settings such as, for example, in Bangladesh (Baliki et al., 2019), Ghana (Yahaya et al., 2018), Rwanda (Nsabuwera et al., 2016), and Tanzania (Larsen and Lilleør, 2014). Combinations of in-kind support with cash transfers also show positive impacts on food security in these settings (e.g., Banerjee et al., 2015).

There are substantial knowledge gaps on the dynamics of the impact of SAI on food security. First, even though most recent impact evaluations of such programs in low and least developed countries show increases in food security, it remains to be investigated if these conclusions can be transferred to conflict-affected settings as agricultural production and food security are significantly challenged by conflicts (Martin-Shields and Stojetz, 2019). Second, there is a vivid discussion on whether direct transfers effectively complement the impact of SAI. Third, other adverse shocks like climatic events, crop pests or health shocks are also known to negatively affect agricultural production and food security, however, insights concerning the impact of these shocks on outcomes of SAI are scarce.

This thesis provides a large-scale assessment of short-run causal effects of a SAI on food security in least developed contexts and zones affected by conflict with evidence from South Sudan. With the help of quasi-experimental household panel data, this thesis will study the impact of an assistance package including a combination of asset transfers, capacity-building methods, and an emergency cash transfer on food security in a peri-urban region in Jubek State, South Sudan. A Difference-in-Differences approach combined to Propensity Score Matching guarantees a high reliability of the assessed impact to be causally attributed to the program. The analysis will differentiate the impact of a one-time cash transfer to flood-affected households from the combined agricultural intervention. Additionally, the analysis will shed light on the role of adverse shocks on the treatment effect.

The analysis reveals that a SAI in a setting highly affected by conflict could shift the respondents' food composition towards a more diverse diet within one year, increasing the Minimum Dietary Diversity Score by approximately 1 point ($p<.01$). Due to the support, beneficiary respondents shifted their diet away from green leafy vegetables towards other vegetables, pulses, nuts, meat products, and eggs. However, sugar intake also increased due to the program. The Food Consumption Score confirms changes in dietary composition through the intervention, but clear improvements are not assessable. Even though, the treatment impact on the Food Insecurity Experience Scale is as good as zero, the likelihood for respondents to experience a whole day without food decreased by 7% points ($p=.08$) through the program. The harvest clearly increased through the program for all crop groups: by 50 kg for cereals and by 19kg-28kg for other main crops, vegetables, and fruits ($p<.01$). SAI combined to a one-time cash transfer are attributed to the treatment impact on MDD, so that it is assumed that the transfer overcompensates for the impact of flood catastrophe. Adverse shocks mainly do not disturb the intervention impact. A severe disease or an accident with the head of the household offset the positive impact in dietary diversity.

The knowledge on how to improve food security in conflict-affected least developed settings provides a key step towards reaching SDG2. This thesis makes the following contributions to the formal literature: First, it indicates that the existing knowledge of SAI concerning the impact on agricultural production from developing settings can be transferred to a least developed context, highly affected by conflict in permanent risk of adverse shocks. However, to apply the existing knowledge on food security outcomes, there is still a dearth of evidence. Second, this thesis presents the implementation recommendation to support farming households with combined agricultural approaches, including in-kind components and cash transfers in situations of emergency. This is particularly valuable for decision-makers to effectively allocate resources, also in other regions of the world with similar challenges.

The second section provides the reader with the relevant economic theory and recent literature of the dynamics between food security, agricultural interventions, and adverse shocks, and supplies important context-specific information. Section three describes the analysis approach including methods and data. The fourth section will focus on the main findings, reinforced by robustness checks. In the fifth section, the findings will be discussed in a wider context. This thesis ends with the conclusion that SAI support small-scale farmers in emergency settings; however, they are no stand-alone solution.

2 Background

2.1 Theoretical Background

2.1.1 Food Security

"Food security exists when all people, at all times, have physical and economic access to sufficient, safe and nutritious food to meet their dietary needs and food preferences for an active and healthy life" (FAO, 1996). Consequently, food security is a multi-dimensional term that can be estimated with a variety of indicators addressing different components. Divergences in assessments might emerge with different measurement approaches (Maxwell et al., 2014).

Smith et al. (2000) identify national food availability and people's inability to access food as the most basic causes of food insecurity. The majority of the food insecure live in underdeveloped settings (Sanchez et al., 2005), where large shares of the population are engaged in agriculture[4]. Small-scale farmers are estimated to comprise half of the global food insecure[5] (ibid.). It appears paradoxical that food producers are food insecure but small-scale farmers in developing countries frequently live in conditions of extreme poverty (ibid.), which challenges the purchasing of additional food. Further, harvest loss and inadequate conservation impede a steady access to adequate nutrition among poor rural households (Shimeles et al., 2018). This enhances the high seasonality of food insecurity (Yahaya et al., 2018). Also, subsistence farming is often too inefficient for self-sufficiency (Sibhatu and Qaim, 2017). Moyo et al. (2015) point out that in Africa the following rea-

[4] In low-income countries, on average 59% of the employment in agriculture in 2019 (World Bank and ILO, n.d.).

[5] The definition of small-scale farming is context specific in the literature. In this work, farmers are defined as smallholders if they operate less than 2 ha of arable land in line with Dalberg (2012).

sons predominately impede the necessary agricultural productivity: weather shocks in combination with high rain dependency, the lack of adaptability of agricultural innovations to the African context, impeding market regulations, poor market infrastructure, low investments, and institutional failures. A further challenge is the increasing rivalry for land and water through population growth (IFAD and UNEP, 2013; Lowder et al., 2016). Jayne et al. (2010) point out that small-scale farmers in Eastern Africa face a decreasing ratio of land cultivated per person engaged in agriculture by 25-50% from 1960 to 1999, in addition to stagnating crop productivity and decreasing donor and state support.

2.1.2 The Impact of Adverse Shocks

The risk of adverse shocks challenges the rural poor in the underdeveloped world permanently (Banerjee and Duflo, 2011). Shocks induce welfare losses through unexpected asset destructions or income losses (Dercon, 2004). Mostly, shocks threaten one source of income, wherefore, farming households in low and least developed countries are particularly vulnerable because they often depend on risky and non-diversified income sources, while welfare states are underdeveloped, insurance mechanisms are barely deployed, and access to formal credit and savings is scarce or quick to exhaust if they exist (Banerjee and Duflo, 2011, Andersen and Cardona, 2013). To ensure the satisfaction of the most basic needs, vulnerable households are often forced to adapt harmful livelihood strategies to cope with a shock. One such strategy is the reduction of consumption (Gao and Bradford, 2018). This might include a decrease in food spending or in consumption of one's own harvest, so that income can be generated through sale. Further, households might react with the sale of household and productive assets to regain liquidity in the short run. Particularly the latter impedes income and productivity in the long run (Doss et al., 2018). For farming households, the sale of livestock, for example, can bridge a gap in the short run but impedes further income generation thus making the household more vulnerable to future disturbances (Deaton, 1991). Consequently, most poor farming households are highly risk averse. To minimizing their risk, they tend to allocate their resources inefficiently (Arias et al., 2019).

Economists distinguish between idiosyncratic shocks, which are experienced at the individual or household level, like health shocks, death or unemployment; and covariate shocks, which affect the broader community like epidemics, climatic shocks, crop and livestock pests, conflicts, or price shocks (Günther and Harttgen, 2009). While the first shock type rather leads to people's inability to access food, the latter can contribute to food supply difficulties on a broader level, as well as to individual access inabilities (Maxwell and Smith, 1992).

In 2021, out of the ten worst-affected countries by food crises, six are experiencing severe protracted conflict (FSIN, 2021). Conflict often has a causal negative impact on agricultural production and food security (Brück and d'Errico, 2019; Martin-Shields and Stojetz, 2019). For example, Arias et al. (2019) point out that farmers in Colombia adapt less efficient production methods when they are exposed to conflict. Verwimp and Muñoz-Mora (2018) show that recently returned internally displaced persons (IDP) in Burundi show lower food expenses and caloric intakes compared to their non-IDP neighbors[6]. George et al. (2020) elaborate that Boko Haram's attacks in Nigeria decrease peoples' dietary diversity. On the contrary, Brück and d'Errico (2019) demonstrate using the distance to the Israelian border as an instrument for conflict exposure in Gaza that Palestinian households' food security is not directly diminished by conflict. Martin-Shields and Stojetz (2019) stress that food insecurity can also enhance conflict. Brück et al. (2015) elaborate with a literature review that rigorous impact assessment is particularly challenging in conflict-affected settings. Furthermore, d'Errico et al. (2021) emphasize with evidence from South Sudan that highly conflict-affected regions receive a lower amount of humanitarian assistance.

Additionally, plenty of literature points out the alarming relationship between further adverse shocks and food security of the poor, accounting for agricultural production or assessing the direct relation. For example, Gao and Bradford (2018) investigated on household consumption behavior in connection to adverse weather shocks in rural Ethiopia using panel data from 1994 to 2009. In years with below average rainfall, the average consumption level per adult was 18.2% lower than in years with above average rainfall. Also, in years, in which the days of extreme heat doubled in comparison to the previous year, real consumption dropped by 3.4%. With logistic regressions, Heltberg and Lund (2009) show that the probability of experiencing food shortages for poor households in rural districts of Pakistan increases with economic, natural, and agricultural shocks. Also, after idiosyncratic health shocks, the households are more likely to face food scarcity. All in all, 32% of the studied adverse shocks resulted in food insecurity.

2.1.3 Pathways to improve Small-Scale Farmers' Food Security

In contrast to adverse shocks, development aid is intended to leverage receivers towards more welfare (Keeley, 2012). Already 40 years ago, Sen (1981) stressed the importance of small-scale farmers in the fight against hunger and also in the present literature, researchers stick to this position (e.g., Carletto et al., 2015; Duncan et al., 2020). Apart from researchers; donors, practitioners, and policy-makers agree on the importance of agricultural-based approaches to improve nutritional outcomes in the global south (Levin et al., 2003). Con-

[6] To control for selection biases, the researchers apply propensity score matching and an instrumental variable model.

sequently, the support of small-scale farmers is on many political agendas in low and least developed settings (Abraham and Pingali, 2020).

A common program objective in rural underdeveloped regions is to improve small farm productivity, which, in turn, enhances the households' welfare reducing poverty, food insecurity, and vulnerability towards adverse shocks (Rahman and Westley, 2001; Sanchez et al., 2005). The large list of further development goals which are potentially approached by SAI embrace better health outcomes (Weiser et al., 2015), self-reliance (Nyikahadzoi et al., 2012), economic development (Timmer, 2002), the empowerment of vulnerable groups (Baliki et al., 2019), social cohesion, environmental sustainability, and overall well-being (Waddington et al., 2014). In conflict-affected settings, subsistence agriculture often proves to be one of or the key coping mechanisms to maintain a nutritional safeguard (Arias et al., 2019; Bageant et al., 2016; Bozzoli and Brück, 2009).

SAI designs can be divided into three mechanisms: enhancement, i.e., to support the established practices; diversification, i.e., to complement the traditional practices; and substitution, i.e., to implement new techniques (Fiorella et al., 2016). To improve food security through SAI in the global south, FAO (2017) gives the following recommendations: to target the most vulnerable, to diversify production, to enhance nutrient-dense crop cultivation, to adapt to the local context, and to include educational measures. More concisely, Nyikahadzoi et al. (2012) stress that the adaption of innovations and technology, input access, know-how, and market information are crucial for small-scale farmers to improve their food security in Sub-Saharan Africa. Taking the conflict-context into account, Bozzoli and Brück (2009) observe in postwar communities in Mozambique that farmland extension and strengthening of established agricultural practices contribute to household welfare and that subsistence and diversified farming enhance better food security outcomes.

There is a pool of literature elaborating the impact of SAI on food security. Berti et al. (2004) show by reviewing thirty farm-related intervention reports that agricultural production mostly increases through SAI but the effect on food security is inconclusive. This is confirmed by a more recent review by Bizikova et al. (2020) including seventy-three published impact assessments on the effect of agricultural interventions on food security. Two-thirds of the publications reveal positive impacts, 23% could not prove a significant effect, while 10% even find a negative impact across various intervention types. Wordofa and Sassi (2020) conclude mostly positive impacts on the consumption, diversity, and security of food in the specific context of Ethiopia from their review of twenty-five nutrition-sensitive agricultural intervention evaluations.

Researchers evaluate a range of SAI approaches. Capacity-building approaches are a commonly implemented instrument. For example, Larsen and Lilleør (2014) show with a Quasi-Difference-in-Differences approach that Farmer Field Schools could decrease the probability for households to face hunger measured by the Household Hunger Scale in Tanzania[7] [8]. Yahaya et al. (2018) point out with an Endogenous Treatment Effects model that training in sustainable agricultural intensification practices could increase household food security as measured by the Household Food Insecurity Access Scale (HFIAS) in Ghana[9]. Most of the capacity-building approaches are implemented combined to other measures. Pan et al. (2018) emphasize with a Regression Discontinuity Design, that female small-scale farmers in Uganda could increase their production through capacity-building combined to seed provision. However, from a similar project combining training to input provision in rural Zambia, Rosenberg et al. (2018) also find increases in nutritious crop production but no significant improvements in dietary diversity for young children and their mothers with a Difference-in-Differences approach. Likewise, from a similar intervention, Margolies (2019) could not detect a substantial impact on HFIAS and the Household Dietary Diversity Score with a Randomized Control Trial (RCT) in Malawi[10].

Another possible agricultural intervention approach is home gardening. The micro-scale homestead production of mainly fruits and vegetables generally complements the households' food intake. Galhena et al. (2013) emphasize a positive relationship between home gardens and food security with a literature review. A gardening program often connects trainings and input provision. For example, Baliki et al., (2019) show with a Difference-in-Differences approach that training on gardening and nutrition complemented by seed provision could lead to at least three-year lasting increases in vegetable production and consumption in Bangladesh. Other researchers confirm positive effects on nutritional outcomes through similar approaches (Olney et al., 2015; Schreinemachers et al., 2015, 2016[11]; Tesfamariam et al., 2018; Marquis et al., 2018). Furthermore, technical support is a pathway to enhance agricultural production towards more food security of small-scale farmers in the global south. For example, Tesfaye et al. (2008) stress that with an Endogenous Treatment Effects model that farming households who participate in a governmental irrigation program, show a higher production and improved food security in Ethiopia.

Rigorous impact evaluations of agricultural interventions on food security are mostly available for peaceful settings. However, it is well-known that agriculture, food security, and development support are highly-affected by conflict (Martin-Shields and Stojetz, 2019; Brück and d'Errico, 2019; d'Errico et al., 2021). The collection of data in these contexts is

[7] The studies by Larsen and Lilleør (2014), Yahaya et al. (2018), Nsabuwera et al. (2016), and Margolies (2019) use scales that assess if and how frequently households face certain food insecurity situations (Ballard et al., 2011; Coates et al., 2007)

[8] The Quasi-Difference-in-Differences approach exploits self-selection patterns of an additional treatment group that enrolled later into the program to predict the selection bias.

[9] The model uses Heckman's two-step approach accounting for the decision for program enrollment to control for a selection bias (Wooldridge, 2012a).

[10] This score accounts for the household intake of different food groups within a certain period.

[11] This is the same project like Baliki et al. (2019) assessing the program impact on the short run.

constrained by security, ethical, and methodological obstacles (Puri et al., 2017). To my knowledge, the only related literature indicates positive impacts of agricultural support in conflict phases on food security and child stunting (Mary et al., 2020; Rao, 2021). Due to the dearth of evidence, it remains to be tested if the existing literature on peaceful settings can be transferred and applied to conflict-affected regions.

One oft-considered expansion of or alternative to SAI are cash transfers. There is a lively debate on which approach is more effective to improve smallholder farmers' food security or if some sort of combination is the best practice. In-kind instruments are often preferred by decision-makers to unconditional transfers because these provide the possibility to set directional impulses (Cunha, 2014), to avoid self-selection of people that are not in need in case of asymmetric information (Blackorby and Donaldson, 1988), and to prevent labor supply adjustments through income effects[12] (Cesarini et al., 2017). On the contrary, equal-valued unconditional cash transfers are rather preferred by recipients (Cunha, 2014), enforce the local market, allow the inclusion of resource poor recipients (Ellis and Maliro, 2013), are better applicable to heterogenous needs, and are generally cheaper to distribute than in-kind transfer (Haushofer and Shapiro, 2017). Therefore, in case of similar impacts, unconditional cash transfers are usually more efficient. One should also differentiate between regular transfers and one, single payment: while one-time transfers are mostly regarded as emergency aid or as an investment stimulus, regular payments are rather assumed to ensure a stable development (Ellis and Maliro, 2013).

Researchers reveal that cash transfers have the potential to improve food security. For example, D'Agostino et al. (2013) evaluate the impact of cash transfer programs introduced between 1992 and 2010 on the prevalence of undernourished people in Sub-Saharan Africa with a synthetic control method[13]. They assess that the prevalence of undernourishment declined after three years through these programs by .6% in Sierra Leone and by 1.3% in Zimbabwe[14]. In Malawi, they assess a reduction by 1.9%, while in Rwanda the prevalence of undernutrition decreased by nearly 4%, in Ethiopia by .9%, and in Mali by .8%[15]. Haushofer and Shapiro (2017) assess with an RCT increases in monthly revenue in agriculture and better food security through unconditional cash transfers paid out to poor households in rural Kenya either in a one-time payment or monthly. Monthly transfers demonstrate a higher impact on food security compared to one time payments (Haushofer and Shapiro, 2017, 2018). Therefore, regular and one-time payments seem to diverge in outcomes.

[12] However, one could argue that also in-kind support leads to income effects.
[13] The transfer mechanisms differ by countries.
[14] The authors classify these two countries as fragile.
[15] These four countries are classified as low-income countries.

Nevertheless, the literature does not indicate clearly if unconditional transfers are more effective than in-kind support or if a combined approach is best practice. With a literature review, Gitter et al. (2017) compare SAI and cash transfers. The researchers stress that both approaches lead to improvements in nutrition outcomes in low and least developed countries, however diverging impact sizes do not indicate clear implications on which approach is better suited. Rather than applying intervention instruments individually, Bizikova et al. (2020) recommend complementing capacity-building approaches by subsidies on inputs, food vouchers or cash transfers. Banerjee et al. (2015) evaluate an approach including multiple components such as grants for productive assets, capacity-building components, temporary cash support, the provision of access to credit, and health information in Ethiopia, Ghana, Honduras, India, Pakistan, and Peru with clustered RCTs. A pooled evaluation stresses that very poor households benefit with significant improvements in food security from this program. Similarly, with a before-after analysis, Nsabuwera et al. (2016) show significant improvements on the HFIAS and the Food Consumption Score (FCS) through a combined intervention, including trainings, access to microcredit, and a cash transfer conditioned to purchase agricultural inputs and livestock in Rwanda[16] [17].

Again, impact dynamics might be very different in conflict-affected and humanitarian emergency settings. D'Errico et al. (2021) stress that in conflict-affected areas cash transfers are less sensitive to access problems compared to other intervention methods. Aurino and Giunti (2021) highlight with their literature review that cash and food transfers have mostly positive impacts on child nutrition in crisis settings. Likewise Tappis and Doocy (2018) concentrate in their literature review explicitly on cash transfers in humanitarian emergencies. Examining five impact assessments, they stress that two studies show that interventions dedicated to maintaining food security could reach their objective through cash transfers (Aker et al., 2011; Schwab et al., 2012), while further two studies assessing the impact on food security improvements could not find a universal inference (Aker, 2013; Hidrobo et al., 2014). In the context of the humanitarian emergency in the Democratic Republic of the Congo, Aker (2017) compares the effectiveness of vouchers and cash transfers with a RCT. The reselling of the vouchers was possible. Providing equivalently-valued vouchers and transfers, the author detected improvements in food consumption and well-being indicators through both approaches, but they could not assess significant differences in outcomes, wherefore the direct transfer is more cost-effective. Schwab (2019) stress that cash transfers in conflict-affected Yemen are invested in agriculture. Brück et al., (2019) provide evidence from nutrition-sensitive support programs in conflict-affected Niger. They stress with a Difference-in-Differences approach that food aid com-

[16] A detailed explanation of Food Consumption Score can be found in section 3.1.3.
[17] The inference of this study prone of time confounders as there was no control group applied.

bined with asset-building support improve anthropometric outcomes in children[18]. In contrast to the other findings, food aid alone cannot improve outcomes. The literature emphasizes that cash and food transfers have the potential to improve household food security, but the evidence on transfers combined to agricultural interventions in fragile settings remains thin.

In conclusion, the literature provides rigorous evidence on a range of SAI approaches on multiple food security outcomes from different settings with diverse assessment approaches. The findings on food security measures are mostly positive. Food and cash transfers also show frequently positive effects on food security outcomes. Still, the literature is inconclusive on whether cash could improve the treatment effect of SAI. Usually, several aid components are applied combined, so that it is hard to disentangle exactly what aid component caused what effect. These insights are in line with Fiorella et al.'s (2016) conclusions that the impact of agricultural interventions on food security is difficult to generalize. Further, most of the available rigorous impact evaluations of SAI on food security are assessed in rather peaceful, low-developed settings or do not consider the conflict dynamics explicitly. Related studies that regard a conflict context directly, mostly do not evaluate SAI. In line with Puri et al. (2017), it is clear, that rigorous assessments of SAI in highly fragile contexts are needed for effective allocation of aid. In section 2.1.2, it is elaborated that adverse shocks challenge agricultural production and food security. It is logical to assume that adverse shocks challenge the effectiveness of SAI to reduce food insecurity, as they hinder aid to function as planned. However, there is no rigorous assessment of the impact of (multiple) adverse shocks in conflict-affected settings on the success of SAI.

Therefore, there are clear knowledge gaps in the literature. First, evidence on the effectiveness of SAI on food security in conflict-affected settings is needed. Second, more knowledge is required on whether cash distribution can enhance the effectiveness of SAI. Third, heterogenous treatment effects of SAI under the influence of other (multiple) adverse shocks is barely explored.

To address these knowledge gaps, the following research questions emerge:

1. Can multi-component SAI contribute to improve food security outcomes in a least developed setting highly affected by conflict?
2. Does a one-time emergency aid transfer combined with the SAI alter food security outcomes?
3. Is the intervention impact sensitive to the occurrence of other adverse shocks?

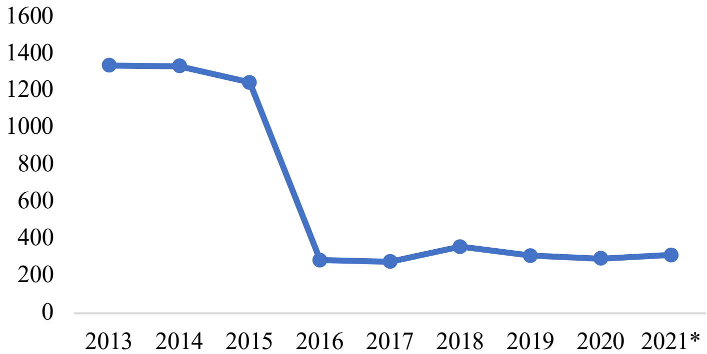
[18] The households are given seedlings for tree to grow lumber as a safety-net and for nature regeneration.

2.2 Humanitarian and Macroeconomic Indicators in South Sudan

The place of study is a peri-urban region in the outskirts of Juba, the capital of South Sudan. The study location belongs to the southern state of Central Equatoria which is crossed by the White Nile.

In 2011, South Sudan obtained independence from Sudan following five decades of recurrent conflict. The lack of governmental representation for the South, perceived discrimination, and inter-tribe conflicts are considered to be main drivers of the conflict (Center for Military History and Social Sciences of the Bundeswehr, 2018). A new civil war was triggered in the nascent republic of South Sudan in 2013 due to conflicts across political power and ethic-based violence (Jok, 2021). The involved parties agreed through a peace agreement to end the civil war in 2018 (Center for Military History and Social Sciences of the Bundeswehr, 2018). The ceasefire has been unstable, and outbreaks of violent actions continued to reoccur even after the peace agreement in 2018. Based on ACLED data (n.d.), the year 2020 has witnessed a sharp increase in violent events in comparison to the previous year by 57%. One of the most noted factors driving the recent increase of violence is around land access and high migration (McMichael, 2016). The South Sudanese people reacted with the establishment of civil self-protection mechanisms and community justice practices (Rhoads and Sutton, 2020). By 2019, the civil war led to approximately 380,000 deaths (FSIN, 2020).

Figure 1: GDP per Capita in US-$



Notes: GDP per capital in US$ in South Sudan (2013-2021); IMF staff estimates; *projection for 2021 (IMF, 2021; own illustration)

During the civil war, the estimated South Sudanese GDP per capita (PPP) indicates a downfall (Figure 1). The estimated GDP settles at around one fourth of its value in 2015 at 278 to 359 USD (IMF, 2021). In 2020, the economy remained unstable. This can be associated mainly with the Covid-19 pandemic and the resulting drop in oil prices[19], floods, as

[19] South Sudan generates around 90% of its export revenue with oil (Macrotrends LLC, 2020; World Bank, 2021a).

well as a locust plague (World Bank, 2021a). With around 11,000 confirmed Covid-19 cases and 120 confirmed deaths, the health impact of Covid-19 seems to be minimal (John Hopkins University, June 2021). However, these numbers are likely to understate the real spread of the virus due to limited testing capabilities and social stigmatization (UNMISS, 2020).

South Sudan is generally prone to climatic shocks like floods and droughts (Ajak, 2018). Recently, severe floods in the second half of 2020 led to 504,000 displaced people and affected over 1 million individuals, primarily in the central and eastern parts of the country, including the study region (OCHA, 2021, 2020). People face extreme food supply shortages through crop loss and infrastructure destruction (Amnesty International, 2020). Moreover, the people were exposed to major health threats through precarious Water, Sanitation, and Hygiene (WASH) conditions and the increased malaria prevalence (OCHA, 2021). Additionally, locust swarms that passed over southern parts of South Sudan in March 2020 and livestock endemics impede the necessary agricultural production (d'Errico et al., 2021b; FAO et al., 2020). Within two years, the average price of food increased significantly in South Sudan. From 15 April 2019 to 15 April 2021, the nominal average commodity price doubled (HDX, 2021). In the state Central Equatoria, the cost of a minimum basket expenditure increased by a factor of 2.7 from June 2019 to June 2021 (CLiMIS South Sudan, 2021). 85% of the people in South Sudan make their living with farming, fishing, and herding (FAO, 2019). Paradoxically, the country showed a net production gap in cereals one-third of its demand in 2019. In Central Equatoria, the net production gap in cereals was even more than half of its demand (CLiMIS South Sudan, n.d.a.).

It becomes clear that a high share of South Sudanese households is exposed to a range of severe threats. In 2021, more than 70% of the South Sudanese depend on humanitarian assistance (World Bank, 2021a). 80% of the population lives in absolute poverty with less than 1.90 USD per day in 2020 (World Bank, 2020). According to the Human Development Index, South Sudan is ranked 185th out of 189 countries in 2019 (United Nations, 2020c). In total, nearly 2 million people are internally displaced in 2020 (World Bank, 2020) and nearly 2.2 million people fled to the neighboring countries (UNHCR, 2020).

The multitude of adverse threats exacerbates the severe condition of food security in the country. In October/November 2020, the Integrated Food Security Phase Classification (IPC)[20] assessed an alarmingly rise of people in food crisis or worse in the whole country (53%, compared to 38% in October/November 2019) (FSIN, 2021). In Central Equatoria the same share doubled to 46% in the same time frame (CLiMIS South Sudan, n.d.b; FSIN,

[20] The IPC is a tool developed by fifteen organizations and inter-governmental institutions to classify the different dimensions of food insecurity and malnutrition among countries. The scale for Acute Food Insecurity is divided into five severity levels: (1) Minimal/None, (2) Stressed, (3) Crisis, (4) Emergency, (5) Catastrophe/Famine (IPC Global Partners, 2019).

2021). In the fourth quarter of 2020, the county Juba was classified as in crisis due to acute food insecurity (IPC Global Partners, 2020).

The peri-urban region of Juba is an insightful study location for the research questions raised in this work. The SAI approach enhances the most prevalent form of livelihood in this region. The recent and protracted conflict in combination to other adverse shocks allows for further study on this understudied issue. Furthermore, the region is in an alarming food security condition, wherefore knowledge on the effectiveness of an intervention helps to allocate resources efficiently and rapidly.

3 Empirical approach

3.1 The Framework

3.1.1 The Intervention

Malteser International (MI) implemented a large, long-term resilience- and nutrition-sensitive program in Juba, South Sudan, with the title *"Strengthened resilience of small-scale farmers and pupils in semi-urban and rural areas of Jubek State"*[21]. In cooperation with local institutions, particularly vulnerable small-scale farming households were assigned to receive support. In total, 1,000 households were assigned to directly benefit from the program. In this way, ~7,000 people benefitted directly from the program. Key targeting criteria was that 55% of the assigned households should been headed by women and 40% of the beneficiary households should have had their own fertile land with access to the river, while the other 60% of the beneficiaries should have depended on rain-fed agriculture.

The program started in June 2019 and ran for a period of 3.5 years until the end of 2022. Table A1 in the appendix illustrates the time frame of the intervention. The participating households were assigned to thirty-five farmer groups through which information and support was distributed. The program consists of different intervention tools, including in-kind assistance like provision of seeds and tools, personal support through monitoring visits, and the expansion and renovation of vegetable and home gardens. Further, the intervention is connected to capacity-building through demonstration gardens and training. One household member attended two training sessions on each of these topics: (1) good agricultural practices, (2) seed selection and breeding, and (3) improved and innovative agricultural systems. Complementary to the agricultural approach, 88% of the beneficiary households

[21] In this work, only the agricultural component of the intervention will be studied.

that were affected severely by the floods in 2020 received an emergency cash transfer between 103 EUR and 307 EUR[22] during the intervention phase. Table A2 in the appendix provides more details of the program components. The assistance aims to improve the participants' access to food, their dietary quality, strengthen their resilience against adverse shocks, such as violent conflict, natural disasters, and price volatility in local markets by enhancing robust agricultural production.

3.1.2 The Study Design

With respect to the existing literature, MI's SAI is expected to increase the production and productivity of targeted smallholder farmers. This is expected to have a positive impact on food security after one year of program implementation. However, as the setting is highly affected by conflicts and other adverse shocks, the impact is possibly sensitive towards these disturbances. Even though, there might emerge different impact sizes adopting different food security measures, the overall impact direction is expected to indicate the same conclusions. Adding a cash transfer to a combined SAI approach is expected to further improve food security outcomes. To verify these assumptions, MI and a research team from Leibniz Institute of Vegetable and Ornamental Crops (IGZ) entered in cooperation in order to conduct a rigorous impact assessment.

The impact evaluation makes use of household survey data. These are collected in a panel data format, so that information on the same entities is collected across time. Panel data are characterized by a small number of periods and a large number of observations (Wooldridge, 2012b). Compared to other data formats like cross-sectional or time-series data, panel data provide a higher degree of causal inference and allow a better tracking of individual behavior (Hsiao, 2007). The present data set consists of two periods, one before the complete implementation of the program and one approximately one year after the program started. The observations are divided into a treatment and a control group. The treatment households received support, while control households did not receive any assistance from MI. The assignment to the groups was non-random. This setup is called a quasi-experiment (Wooldridge, 2012b). In contrast to RCT, which builds on a random assignment to the groups, this experimental design is often target to selection bias. Therefore, RCT show a stronger internal validity. Moreover, RCT assess the intervention impact for the whole population, while quasi-experiments capture the intervention impact on the beneficiary group only (White and Raitzer, 2017). However, RCT is not feasible for this study because it is ethically questionable to randomize support in such a severe setting and the procedure requires higher expenditures (Cameron and Trivedi, 2005a).

[22] The exchange rate for South Sudanese Pesos from 12.08.2020 is used (153,5) (CurrencyRate, 2021). The received amount is relative to household size.

Local enumerators conducted the survey in field. The enumerators received training sessions conducted by the research team from IGZ before each data collection wave. The first training block was conducted on-site in Juba, including one field day. With the enumerators' feedback from the trainings, the questionnaire was reviewed again. The second session was conducted remotely because of the travel restrictions due to the Covid-19 pandemic. The research team recorded video sequences, which the enumerators watched in groups, so that they had the chance to discuss and practice. The research team was available remotely in the event any doubts were to have arisen. The trainings were intended to communicate the study aims and sensibilize the enumerators for field work practices.

The enumerators conducted the survey through face-to-face interviews in the respondents' native language (mostly Arabic). The person responsible for the participation in the intervention was supposed to be interviewed but was also replaceable with another adult household member, who understands the intervention if the corresponding person was not available. The first wave was collected between November and December 2019 and the second wave between November 2020 and January 2021[23]. Both waves were collected in the dry season during the harvesting time of the main food crops.

The survey questionnaire was elaborated with KoBoToolbox. This survey tool allows for offline data collection so that intermittent internet connections do not cause data loss. The coordinating team in South Sudan uploaded the data daily. The research team had continuous access to the data and could monitor the progress of the data collection remotely. The survey tool allows a validation logic to exclude unrealistic answers and typing mistakes. Hence, measurement errors could be restricted as the enumerators had the chance to reexplain questions in case of invalid responses. Therefore, the survey ensures high quality of data. The questionnaire was elaborated in English and translated orally on-site by the enumerators. The surveys in both waves are equally constructed.

3.1.3 Variables

In the questionnaire, different outcome variables are included to measure the impact of the treatment multidimensionally. First, three approaches to measure food security are included to capture a broad perspective on access to, quality, and quantity of food. Second, production data will be used to measure the productivity increase through treatment.

The **Food Insecurity Experience Scale (FIES)** is an indicator developed by FAO that assesses the respondents' self-reported experience of food insecurity based on constraints on their ability to access adequate nutrition (Cafiero et al., 2018). It consists of a set of

[23] Even though, the intervention already started before the baseline, the main production affecting components are subject to a natural impact lag through production cycles and, therefore, do not influence the baseline results.

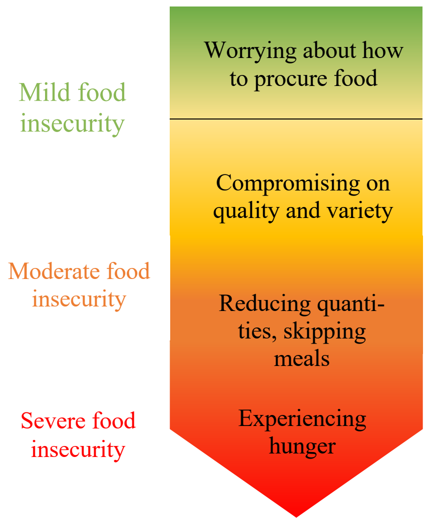
eight short *"yes"* or *"no"* questions each referring to a different experience and insecurity level from being worried about nutrition to not having access to food, see Table 1. Responses to these questions are usually evaluated with the Rasch Model and classified into mild, moderate, and severe food insecurity, see Figure 2[24]. Aware of the recommendation by FAO not to evaluate this score by the occurrence of the separate situations (FAO, n.d.), this will indeed be done in this analysis to assess small changes through treatment. An inverse FIES score from 1-9 is constructed, accounting for the number of *"no"*-responses[25]. In comparison to other measures which focus generally on food quantity or quality, FIES also captures psychological indicators like anxiety and uncertainty of food insecurity (Cafiero et al., 2018).

Table 1: Food Insecurity Experience Scale (FIES)

During the last 12 months, was there a time when, because of lack of money or other resources:
1. You were worried you would not have enough food to eat? (worried)
2. You were unable to eat healthy and nutritious food? (healthy)
3. You ate only a few kinds of foods? (fewfood)
4. You had to skip a meal? (skipped)
5. You ate less than you thought you should? (ateless)
6. Your household ran out of food? (runout)
7. You were hungry but did not eat? (hungry)
8. You went without eating for a whole day? (whlday)

Notes: Adapted from Cafiero et al. (2018)
Acronyms in parenthesis

Figure 2: FIES Classification



Notes: Adapted from Cafiero et al. (2018), own illustration

The **Food Consumption Score (FCS)**, developed by WFP, indicates the frequency of the consumption of different food categories by a household. It, therefore, reflects dietary diversity (WFP, 2008). Household respondents are asked about how many days the household consumed food out of a certain group in the last seven days. The following categories

[24] The Rasch Model measures unobservable traits like food insecurity. By classifying the difficulty to not experience a certain situation depending on the sample average, it estimates the probability that a household with a certain level of threat to food insecurity states to have experienced this situation with a logistic function (Cafiero et al., 2018).
[25] If a household did not experience any of these situations, the score equals nine. It is constructed inversely, to comply with the overall indicator logic that higher scores imply better outcomes.

with the corresponding weights referring to the relative nutritive importance in parentheses are included[26]: *starches (2), pulses (3), vegetables (1), fruits (1), meat/fish/eggs (4), milk/dairy (4), fats (.5), sugar (.5), condiments (0).* The FCS is derived from the sum of the weighted category values. The overall score takes a value from 0 to 112. It is classified into *poor* (0-21), *borderline* (21.5-35), and *acceptable* (>35) food consumption[27] (WFP, 2008). In the analysis, the score is included instead of the classification to account for small changes in food consumption also within the three overall categories.

Minimum Dietary Diversity (MDD) also measures dietary diversity assessing whether individuals ate certain food groups on the previous day. The indicator assesses originally whether an individual consumed food from at least five out of ten food groups. Reaching or exceeding this threshold indicates higher micronutrient adequacy (FAO and FHI, 2016). The food categories include: *grains/white roots/tubers/plantains, pulses, nuts/seeds, diary, meat/poultry/fish, eggs, dark green leafy vegetables, other vitamin A-rich fruits and vegetables, other vegetables, other fruits.* In this analysis, this measure is modified from a dummy variable to a score from 0 to 10 indicating the exact amount of food groups consumed. Again, with this modification, the variation in the indicator can be exploited more accurately, accounting for small differences as well. Because of day-to-day variability of individual nutrition, this indicator is only adequate at the population-level assessing differences in nutrition diversity between groups (ibid.).

In comparison, these indicators cover different dimensions of food insecurity and will lead to broad insights into the impact dynamics. FIES and MDD are asked on the individual level, while FCS covers the household level. FIES captures access to food in different stages of severity. FCS and MDD cover dietary diversity in different time frames. Particularly, these two indicators are expected to show a high correlation.

Agricultural production is used as an additional outcome to assess the program impact. Agricultural production is divided into four groups: Cereals, other main crops, vegetables, and fruits. These crops are divided into different sub-types. In the questionnaire, participants were asked to state details on their maximum three main cultivated plants per group. Among other variables, harvest in kilogram was collected. For this variable, I replaced the outliers in the highest 5% by the 95%-threshold. Households that stated to have not cultivated any crops in one group are computed as zero. Then, I generated the sum of the harvest per group. These four values are included as an outcome to estimate the changes in agricultural productivity over time. Afterwards, the harvest variables will be used as coef-

[26] In the questionnaire, the categories are split into more components to address the typical diet. The evaluation uses the combined score categories, created by merging the subcategories (WFP, 2008).
[27] This threshold is used in countries, in which oil and sugar is not eaten on daily basis (WFP, 2008).

ficients in regressions on the food security to learn about the association between production and food security.

Furthermore, **food spending** will be used as an outcome. The variable is generated as a sum of average weekly expenditure within the past twelve months. The sum includes the following food groups: cereals, legumes, vegetables, fruits, meat, chicken, and milk.

Adverse shocks are of special interest to explore the dynamics of treatment. The survey provides insights on the individual experience of adverse shocks. The respondents were asked whether they experienced the following types of shocks in the past twelve months: *drought, flood, unusual high cost of agricultural inputs, high level of crop pests and livestock diseases, conflict/violence, serious illness/accident of the household head, theft.* If the respondents stated to have experienced a certain shock, they are then asked to rank the impact of the shock by *"not severe"*, *"slightly severe"*, *"highly severe"*, and *"alarmingly severe"*.

According to OCHA (2020), **floods** were overwhelmingly severe in Juba in 2020. For this reason, objective flood data are informative. This shock is of special interest, because highly flood-affected treatment households received additional cash support. Malteser International provides information on the inundations in 2020. QGIS allows to match the baseline observations with the inundated parts (see Figure A1 in the Appendix). However, the data only allows the matching of the floods to the interview location. 7% of the households are matched. The GPS data cannot cover all affected sites, like housing, markets, schools, hospitals, or agricultural plots. Therefore, these data do not provide enough information to be included in this analysis.

Lastly, data on a range of socio-economic characteristics of the respondent, the household head, and the households are collected. These include the gender, age, literacy, occupation, and the marital status of the respondent and the household head. The number of household members was assessed. Further, it was asked if the household has an IDP or returnee status and if they received aid from other organizations during the survey period.

3.1.4 The Sampling

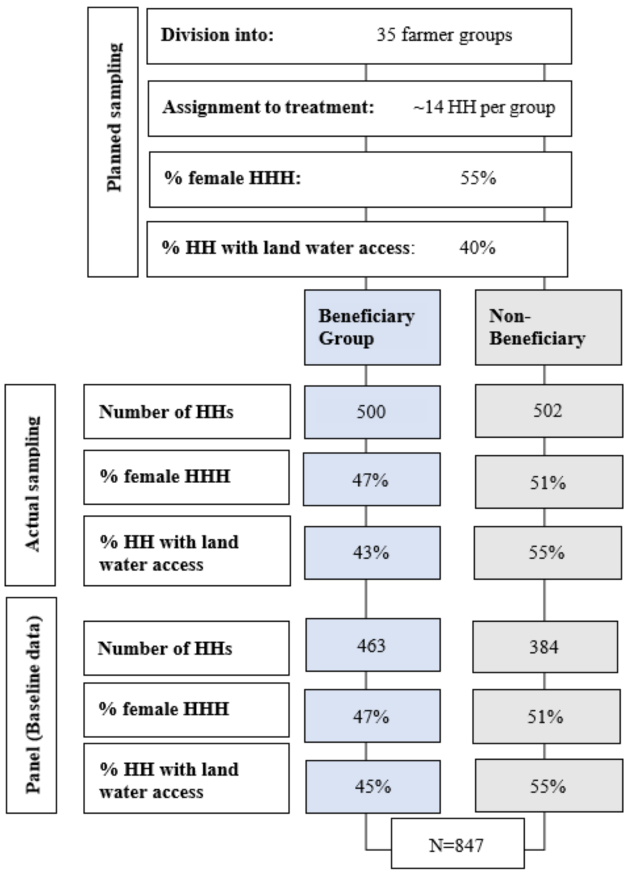
Figure 3 illustrates the sampling procedure. All in all, the objective was to sample 1,000 households, 500 per group, to have enough power for robust results[28]. The treatment group observations were selected randomly from the full set of beneficiaries (N=1,000) with respect to proportional farmer group representation. Hence, around fourteen households per

[28] In this project, no sample size calculation was conducted beforehand. The method helps to determine an efficient number of observations to detect effects with a certain precision without wasting resources on large samples (Columb and Stevens, 2008).

farmer group should be surveyed. To represent the full beneficiary group, the criteria for assignment to treatment should be reflected in the sample. Hence, 55% of sampled households should be female-headed and 40% of the have water access at their agricultural plot. The research team assigned a replacement household for each assigned beneficiary household, which shares the same farmer group and as many key characteristics possible. The replacement households were interviewed if the assigned beneficiary household was not available at the baseline. Control households were assigned on-site. To counteract selection bias, the enumerators were instructed to assign a control sample that contained equal proportions of the key sampling characteristics. The enumerators collected phone numbers to retrack the households for the second wave.

For the second data wave, the same households were interviewed again. The second data collection posed major complications. First, the Covid-19 pandemic challenged the conducting of the face-to-face interviews. Security measures such as wearing masks, keeping distance, and not entering buildings were adhered to. Second, primarily due to the flooding, a large share of participants was displaced. The enumerators tracked a large fraction of these households through calls, asking neighbors, and revisiting old residence locations.

Figure 3: Sampling



Notes: HH=Household, HHH= Household head

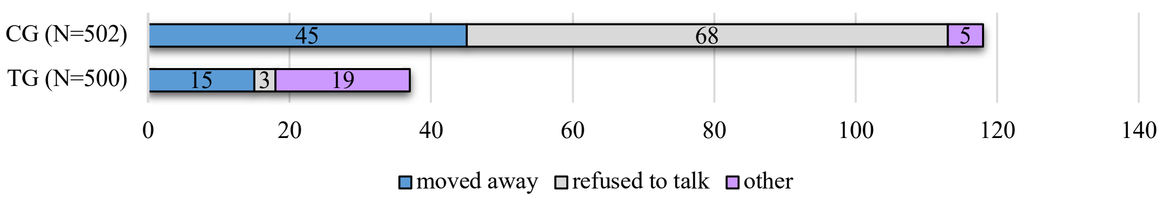
Figure 3 indicates that the enumerators sampled around 500 households for each group in the first wave. The share of female headed households is 47% in the treatment group and 51% in the control group at the baseline. 43% of the sampled households from the treatment group have land water access, while 55% have land access to water in the control group. The enumerators were able to interview 463 households from the treatment group and 384 households from the control group in both waves, reaching a total of 847 panel observations. The sampling characteristics are similar in the first sampled observations and the actual panel and approximately meet the sampling objective.

3.1.5 Attrition

A threat to model validity in quasi-experimental data is systematic attrition. If attrition can be attributed to assignment to treatment, i.e.., if attritted households in the treatment group differ systematically from the attritted households in the control group and these traits are associated with the outcome, attrition causes bias. If characteristics differ systematically between attritted and non-attritted households (independent of household group), external validity is limited.

Even though, the idea was to interview all 1,002 households that participated at the baseline, 155 (15%) of these households are not included into the panel. Figure 4 gives an overview of reasons for attrition by household group. It is evident that attrition was higher with 24% in the control group compared to 7% of the treatment group. There were two main reasons for attrition: no consent to the interview or the household moved away. With 14% of the control group and only 1% of the treatment group, clearly more control households refused to talk to the enumerators. This is reasonable because control households do not benefit directly from cooperating. The treatment households were generally easier to track because they were in constant contact with the MI team on-site. Households moving was also more prevalent within the control group with 9% compared to 3% in the treatment group. Eleven observations from the second wave could not be matched to the baseline[29]. Further, ten observations from the treatment group are excluded because the respondents stated not to have received any support, these are included in *"other"*.

Figure 4: Reason for Attrition by Household Group



[29] Information not displayed. This is around 1% of the full baseline sample. As this is too small to have an impact on the conclusions, these observations are not further regarded.

Table 2 compares socio-economic characteristics of attritted and non-attritted households by household groups. The features of the attritted and non-attritted households appear mostly balanced. However, there are some significant differences. In the control group, household heads of the attritted are on average 3.5 years younger and they have on average one additional household member compared to the non-attritted. The treatment group shows significant differences between groups in literacy of the household head, and their occupation. However, in this group, attrition is with only 7% very low, therefore, hypothesis testing is likely to be misleading. For this reason, a comparison of attritted households between the control and treatment group is also misleading. Table A3 in the appendix confirms that attrition is not concentrated on a certain neighborhood, even though the attrition shares differ substantially between the neighborhoods[30].

Table 2: Attrition Comparison of Baseline Characteristics

	CG			TG		
	Non-attr.	Attr.	p-val*	Non-attr.	Attr.	p-val*
N	384	118		463	37	
HHH characteristics						
% female	50.5	51.7	.824	47.3	45.9	.876
Mean age (years)	43.5	39.9	.009	46.3	49.0	.292
% literate	32.0	35.6	.480	33.3	18.9	.043
% single/widowed	18.5	17.8	.676	19.2	27.0	.312
% farmer	85.2	89.0	.264	95.1	100.0	<.001
HH characteristics						
%IDP/returnee	9.9	11.0	.732	6.5	2.7	.204
Mean Size	8.2	9.3	.058	8.7	7.8	.123
Mean food spending (EUR/week)	50.4	53.1	.626	41.0	36.0	.354
% received other support	18.0	12.7	.151	20.5	32.4	.145

Notes: *p-value with unequal variances (welch), CG=Control Group, TG=Treatment Group, HHH=Household head, HH=household, attr. = attritted

The attrition analysis does not detect any alarming patterns, so that attrition is unlikely to bias the analysis severely. The analysis will cover the restricted panel, only including observations of households that were interviewed twice because no substantial power loss is expected through attrition and biases are not expected.

[30] The neighborhoods are defined based on secondary data on eleven community clusters provided by the local MI team. These data are accessible for 90% of the beneficiary households. In combination with the GPS data, the remaining 10% of the beneficiary households are matched to the community clusters using QGIS. Likewise, the control households are aggregated to corresponding clusters by spatial intersection. The clusters are split into thirty-one subclusters respecting household agglomeration, rivers, and blocks. Because of mobility, baseline locations are used for clustering.

3.1.6 Ethical constraints

Evidence of the impact of small-scale agricultural interventions is scarce in contexts comparable to the South Sudanese case. However, it is crucial to identify what works and what does not. Therefore, rigorous evaluations are necessary to justify the spending of public funds, to set impulses to expand and replicate successful intervention designs, and to extend the knowledge on impact dynamics. To prove causal impacts in this context, it is crucial to assess the development of the outcomes in absence of treatment with a control group. However, the use of a control group is ethically controversial, as the control group does not benefit from the program during the study period, even though people in the control group are clearly in need regarding the dire local situation. To ensure ethic responsibility, the control group should not have disadvantages due to the study. The only "cost" the control group has is the effort for the interview. Malteser International provided a bar of soap for the control group participants after the second interview to appreciate the effort. Furthermore, in case of an expansion or replications of Malteser International's model, the control group is likely to benefit in future as well. And in case of less favorable conclusions, the program design is likely to be adjusted to become more efficient, so that both groups can receive better aid in the future. In this case, the control group does not run the risk of suffering from ineffective or possibly harmful interventions.

The team is committed to the "do no harm" approach. The survey motives, the procedure, and the data usage were communicated with the respondents beforehand. Personal data are kept confidentially and are only used for matching and tracking the households. All publications will be fully anonymized. The participation requires verbal consent from the respondents and is voluntary. The respondents could have revoked their consent at any point of time during the interview. Beneficiary households are not denied further support in case of no participation in the survey. Likewise, support from other organization is not withheld from any of the participating households because of this study. Information on whether the households received support from other organizations was collected to control for heterogenous impacts. The enumerators ensured a pandemic-conform survey in the second wave. The survey received approval from all relevant authorities.

3.2 The Model

3.2.1 The Difference-in-Differences (DiD) Approach

The data allow to compare the outcomes for all panel household before and after treatment. The control group allows to observe the general change in outcomes over time in absence of treatment. Assuming that there are common trends between the control and treatment group, it is possible to estimate the counterfactual of the treatment group, i.e., the hypothetical outcome of the beneficiary group in the second period in absence of treatment (Cameron and Trivedi, 2005b; Wooldridge, 2012c). Taking the difference between the outcome of the treatment group in the second period and its counterfactual delivers the average treatment effect of the treated (ATT)[31]. This is the Difference-in-Differences (DiD) approach. Equation (1) shows the calculation of the DiD estimator $\hat{\delta}$, which assesses the difference between the average outcome $(\bar{y})$ of the treatment group (TG) and the control group (CG) in the difference in outcomes after (t=1) and before (t=0) the intervention. Thus, $\hat{\delta}$ measures the ATT (Wooldridge, 2012c).

(1) $\hat{\delta} = [\bar{y}_{t=1} - \bar{y}_{t=0} \mid TG] - [\bar{y}_{t=1} - \bar{y}_{t=0} \mid CG]$

Implementing the DiD estimator $\hat{\delta}$ into a pooled ordinary least squares (OLS) regression estimation leads to equation (2). This regression requires the standard OLS assumptions (Brüderl, 2010).

(2) $y_{it} = \beta_0 + \beta_1 * wave_t + \beta_2 * treat_i + \delta * (wave_t * treat_i) + u_{it}$

The subscript i indicates household variation, while t implies time variation. y_{it} is the household-specific outcome in a certain period. β_0 is the intercept. β_1, β_2, and δ are coefficient for dummy variables. $wave_t$ equals 1 for observations of the second wave and indicates changes in outcomes across time, independent of treatment. $treat_i$ is 1 if the household belongs to the treatment group and captures group-fixed difference at the baseline. The DiD estimator δ measure the ATT and is the main estimate of interest. $(wave_t * treat_i)$ is only equal to 1 for the treatment group in the second period. The household and time specific deviation of the estimated outcome from the true outcome is given by u_{it} (Wooldridge, 2012c).

The method is robust in non-random assignment settings. The baseline difference (β_2) accounts for all time-invariant unobserved differences between the control and treatment group (Fredriksson and Oliveira, 2019). For example, as MI selected particularly vulnerable households for treatment, the households in the control group might systematically

[31] That the attrition in the treatment group is very low is a necessary condition for the estimation of the ATT. Sampling was randomized within a full set of beneficiaries to approximate the effect for all beneficiaries: a high attrition of the treatment group would decrease the degree of randomization, impeding the approximation of the true ATT.

present different outcomes than the treatment group in absence of treatment. The development of the control group captures changes in the outcome over time that are not due to the treatment but are expected to affect both groups equally, such as general improvement of food security due to a good harvest (ibid).

3.2.2 Model Assumptions

The first and most crucial assumption for the DiD model is the common trend: In absence of treatment, both groups follow the same development. Every influence besides treatment affects both groups in the same way, so that the only difference in development must be due to the support (Angrist and Pischke, 2015). The difference between the groups is, therefore, assumed to be constant (Fredriksson and Oliveira, 2019). If this assumption holds, the DiD estimator indicates the unbiased causal effect of treatment (Meyer, 1995). The common trend cannot be proven as counterfactual data are only hypothetical, but there are indicators which support the reliability of this assumption[32].

The second assumption requires comparability of the two groups. This assumption is a good indicator of the plausibility of the common trend (Fredriksson and Oliveira, 2019). In this case, the control and treatment group are assigned based on the similar key characteristics and the respondents of both groups live very close to each other. Balance tests will evaluate this assumption. Furthermore, Propensity Score Matching techniques are valuable tools to increase the group comparability[33].

The third assumption is the stable unit treatment value assumption. Treatment should not lead to any spillovers or other externalities for the control group (Duflo et al., 2007). However, minor spillovers are possible. First, if the program is effective, control households will indirectly benefit from the program because it will increase the availability of food on the local market. The production through the program might increase the supply, while the participants' demand might decrease. This would attenuate the assessed ATT. Because of the small program scope, the market effect is expected to be minor in the short run. In the longer run, control households are likely to benefit indirectly from the program through overall regional development. Second, control households might adopt agricultural practices from the beneficiary group. The only component of the treatment allowing copying is training. Third, the cash component of the support was not distributed uniformly among the beneficiary households. The main analysis will assess the composite treatment effect, ignoring different treatment intensities. Then, subsampling will be deployed to account for treatment group heterogeneity. Fourth, other NGOs operate in the survey area which might

[32] A multi-period comparison between the control and the treatment group before the intervention would help to classify the robustness of the common trend assumption. The required data are not available in this setting.
[33] The method is explained in detail in section 3.2.3.

affect the outcomes. It is likely, that other relief aid programs affect the relevant outcomes. The possible bias is reduced by including this variable into a propensity score approach.

The fourth model assumption implies that there must not be any changes in group composition. Households should not be able switch groups (Fredriksson and Oliveira, 2019). As the data have a panel structure and no further households are assigned to treatment after the baseline, control households cannot join the treatment group. Treatment households that withdrew from the intervention or state that they did not receive the support, are not included into the evaluation. Therefore, a switch between the groups is impossible. However, withdrawing from the study during the intervention can disturb the common trend. Households that drop out of the program might be less motivated. Another concern is that very poor households are more likely to get displaced, while treatment is expected to lower this risk. Therefore, if a significant fraction of the vulnerable control households drops out, the model will suffer from a downward bias[34].

The fifth assumption emphasizes that being part of an intervention group must not have an impact on the baseline outcomes (Fredriksson and Oliveira, 2019). The outcomes are mainly food security indicators, so it can be assumed that the households will try to reach the best real outcomes possible, independent of being assigned to an agricultural intervention. Consequently, real food security outcomes are unlikely to be disturbed by the knowledge of (not) being assigned to a treatment. However, as the survey is self-reported and consists to a large extent of rather subjective indicators, households might alter their outcomes to elicit (more) assistance or because they are ashamed of their situation. Systematic under- or overreporting will bias the real effect, especially if the extent differs between the groups. As the correctness of non-objective responses cannot be observed and both groups have this incentive in both waves, it is assumed that the responses are valid. Objective data would help to classify the reliability. Again, that the intervention already started before baseline is assumed to have no impacts on the baseline outcomes.

3.2.3 Model Extensions

88% of the beneficiary households that were severely flood-affected received a cash transfer, additionally to the treatment. This leads to heterogeneities within the treatment group. To distinguish between the impact of the cash transfer and the agricultural support, the DiD analysis is run on subsamples. First, the analysis is run only on the beneficiaries using the cash transfer as the treatment coefficient. Treatment households which did not receive the transfer are used as a control group. The DiD estimator will indicate the impact of the cash transfer attenuated by the flood effect. Second, the treatment group is split up into

[34] Again, the attrition monitoring did not detect problematic patterns in dropouts. Propensity score matching will contribute to improved comparability.

those, who received cash and those who did not. Each treatment subgroup is matched to the full control group. Like this, the ATT of the agricultural intervention, excluding the transfer, and the ATT of the composite intervention are compared. Third, using food spending as an outcome, I will be able to elaborate whether the cash transfer was spent on food or rather on other goods. Assuming that the cash transfer was not invested in agriculture, impacts on production depend on the SAI.

The assignment to treatment was non-random, wherefore the model is in risk of a selection bias (Duflo et al., 2007). The subsampling might lead to further imbalances. If differences between the group are time-invariant and do not induce trend divergences, the baseline difference of the DiD model controls for these. However, there might be disparities that lead to different trends, e.g., a substantially higher share of one group gains its livelihood with agriculture, so that these households are more prone to climatic events. This could disrupt the common trend.

Propensity Score Matching (PSM) is a valuable tool to increase the group comparability and to counteract a selection bias. This method uses a vector of observable household characteristics to estimate the probability to participate in the intervention, this is the propensity score. The score is derived with a probit regression. With this score, similar households can be matched between the control and treatment group, so that the balance between the groups increases and the counterfactual can be estimated more precisely (Cameron and Trivedi, 2005b). The first assumption for PSM to be robust is that conditional on the included control variables, the assignment to treatment is as good as randomized (Abadie and Imbens, 2006). Second, all observations should have a positive probability to be assigned to either of the two groups, so these need to be sufficiently similar to be matched (Heinrich et al., 2010). The kernel-based matching algorithm is applied in this thesis. For each observation of the treatment group, the full set of control observations is weighted according to the distance to the certain treatment observation (ibid.). This is also applied in the subsample analysis.

For the PSM, included variables should be strongly related to program participation and to the outcomes (Cuong, 2012). PSM uses the baseline values of the variables, so that the included covariates should be time-invariant (Heinrich et al., 2010). The more deterministic features for the assignment to treatment are included in the estimation of the propensity score, the more the setup approaches a true randomization (Rosenbaum and Rubin, 1983). Table 3a shows correlations of potential covariates for the PSM approach with the main outcome variables. As 82% of the respondents are also the head of household, only the characteristics of the household head are included because their characteristics are assumed to be more informative in terms of the households' welfare[35]. The age and the gender of the respondents are included as these characteristics are assumed to influence the responses substantially. For *"support from other organization",* the correlations are assessed with endline values. This variable is also rather time-variable. It is reasonable to include the covariate like this to explicitly match for assistance during the intervention phase. Mean weekly food spending, land size, and land access to water are assumed to be constant in absence of the program. All these variables are considered to have an association with the outcomes as they are probably good proxies for the households' economic wellbeing. Table 3a emphasizes that all variables have a significant relationship to at least one of the main outcomes. Table 3b indicates the group balance in these characteristics. A range of variables shows significant differences between the groups. The blue-shaded variables are assigned as PSM covariates because they differ significantly between the two groups and are strongly associates with at least one of the outcomes.

Table 3a: Correlations of potential Covariates for the PSM to main Outcomes (Baseline Values)				**Table 3b: Baseline Group Balance of potential Covariates for PSM**		
	Inverse FIES	FCS	MDD	CG	TG	Significance
Female resp	-.22	-.56**	-.03	74.22%	54.86%	***
Female HHH	-.10	1.31*	-.10	50.52%	47.30%	
Age resp (years)	-.01	-.07**	-.02***	40.17	45.16	***
Age HHH (years)	-.01*	-.05*	-.02***	43.54	46.27	***
HHH farmer	.21	5.11***	-.73***	85.16%	95.03%	***
HHH literate	.81***	3.19***	.55***	32.03%	33.26%	
HHH married	.94***	1.96**	.36**	80.47%	80.78%	
HH size	-.03	.52***	.01	8.22	8.66	
IDP/returnee HH	-1.21***	.15	-.09	9.9%	6.5%	*
Mean weekly food spending (€)	.00	.06***	.01***	50.38	41.00	***
Land size (ha)	-.11	3.60***	-.08	.75	.70	
Land access water	.97***	3.76***	.64***	55.47%	44.71%	***
Support from other organization (t=2)	-.31	1.68	-.56***	6.25%	19.87%	***
Severe flood experience (t=2)	-.23	.59	-.32**	61.7%	64.4%	

Notes: Support from other organization and severe flood experience are assessed in endline. Resp=respondent, HH=Household, HHH= Household head, Significance for 3b according to t-tests with unequal variance (welch), * $p < .1$, ** $p < .05$, *** $p < .01$

[35] 98% of the observations had the same respondent in both waves, therefore the respondents' characteristics are seen as time-invariant.

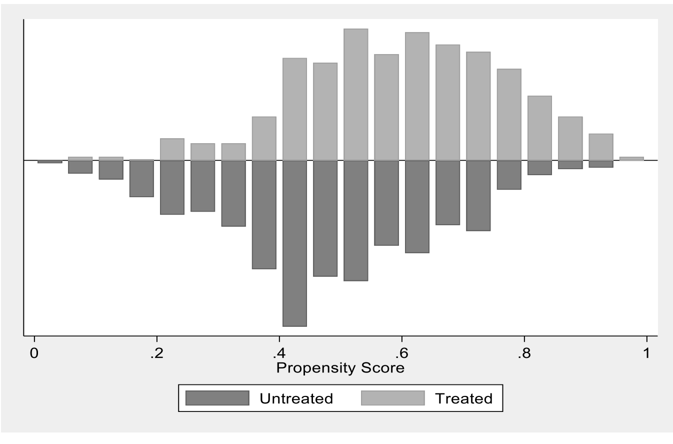
PSM assumption 1 is unlikely to hold, regarding unobservable variables that affect participation in treatment and the outcome such as motivation or vulnerability as mentioned above. These unobservable variables cannot be included into the PSM directly. Indirectly, the PSM approach accounts for household wealth and wellbeing indicators, so that the approach approximates the model closer to randomization. Table 4 emphasizes that after PSM, the former unbalanced covariates are balanced between the groups. Figure 5 shows the distribution of the propensity score for the treatment (*"treated"*) and control (*"untreated"*) group. The figure highlights with a clear overlap that common support is given, i.e., there are matching control households for nearly all treatment households. Therefore, the second assumption holds. For the subsampling, the PSM evaluations are assessed in the Appendix (see Tables A4-A6 and Figures A3-A5 in the appendix).

Table 4: Baseline Group Balance after PSM

	CG	TG	p-value
Female resp	56.3%	54.9%	0.737
Age Resp (years)	46.4	45.2	0.349
Age HHH (years)	46.5	46.3	0.844
HHH farmer	93.8%	95.0%	0.407
IDP/Refugee HH	7.9%	6.5%	0.499
Mean weekly food spending (€)	41.8	41.0	0.817
Land access water	41.8%	44.7%	0.474
Support from other organization (t=2)	16.5%	19.9%	0.375

Notes: P-values from t-tests with unequal variances (welch), Resp= Respondent, HH= Household, HHH=Household head, CG=Control group, TG= Treatment group

Figure 5: Distribution of Propensity Score across Groups



Serial correlation leads to an overestimation of the test statistics (Bertrand et al., 2004). If observations are correlated to each other, the standard errors (SE) are not independent and identically distributed. There is less variation than in an independent sample. This violates the OLS assumption of zero conditional mean (Cameron and Trivedi, 2005c). Clustered SE allow correlation within clusters, assuming the absence of correlation between clusters (Cameron and Miller, 2015). Further, clustered SE correct for heteroskedasticity. Clustering does not change the value of the coefficients but increases the SE of the regressors: the narrower the clusters, the larger the SE (Wooldridge, 2012d). A small number of clusters can cause misleading conclusions through biased SE (Angrist and Pischke, 2008). The common thresholds for the number of clusters is around forty (Esarey and Menger, 2019; MacKinnon and Webb, 2017). It is crucial to select clusters that are congruent with the sampling logic (Abadie et al., 2017).

There are two issues that need to be considered in this analysis. First, as all households are surveyed in both periods, the observations from the same household are not independent from each other, they suffer from first-order serial correlation. Therefore, the SE will be clustered on the household level. Second, household characteristics are likely to correlate within certain groups, e.g., neighborhoods (Cameron and Miller, 2015; Wooldridge, 2012d). Clustering on a group level is reasonable, if groups with a high within-group correlation are detected (Fredriksson and Oliveira, 2019). In this work, neighborhood clusters are considerable[36]. Food security is likely to be more similar within clusters with production and market similarities, comparable adverse shock experiences and similar to water infrastructure. This approach is not applied because thirty-one clusters are below the common threshold, the assignment to the study was independent of clusters, and the cluster size is very unbalanced.

Fixed Effects (FE) is an alternative to the DiD approach. This model makes use of within-variation (Cameron and Trivedi, 2005c). Regressions are set up for both waves separately, subtracted from each other, and then the differences are regressed by standard OLS (ibid.). SE differ between the two models (Imai and Kim, 2020). In a balanced panel, there is no numerical difference in the ATT between the DiD with pooled OLS setup and the fixed effects setup, however the models make use of different assumptions[37] (Cameron and Trivedi, 2005c). In case of attrition, pooled OLS and FE estimates differ because FE works with a balanced set, so that observations that are not assessed in both periods drop out. Pooled OLS allows the inclusion of attritted observations. This would be inconsistent in case of systematic attrition. Pooled OLS is more efficient than FE because it makes use of

[36] Figure A2 in the Appendix illustrates the possible neighborhood clustering approach.
[37] The most crucial FE assumptions are that there must be zero-conditional mean between the time-varying error and the regressors and that there must be over-time variation in the regressors (Cameron and Trivedi, 2005b; Wooldridge, 2012b).

more data[38]. Consequently, if attrition is random, pooled OLS is more appropriate[39] (Lechner et al., 2015). Both models allow a consistent causal interpretation if the assumptions hold. Because both models lead to the same coefficients in this two-period balanced setting, I will only include the pooled OLS model with DiD in the main analysis. For the evaluation of the impact of agricultural production on food security, FE is applied.

The pooled OLS model allows an intuitive interpretation of a triple-difference model, which can control for heterogenous treatment effects (Wooldridge, 2009). In this context, especially adverse shocks are expected to disturb the treatment impact on food security and lead to a heterogenous treatment effect because shock occurrence is time varying. On the one hand, if the intervention was successful, beneficiary households might be more resilient towards adverse shocks than control households. On the other hand, adverse shocks might repress the positive impact of the intervention. The extended model allows a second classification into two groups, in this case: severely shock-affected and not-severely-affected. Dummy variables are generated for each of the seven assessed adverse shocks. If a respondent states to have experienced a certain shock "*very severe*" or *"alarmingly severe"* during the interventions phase, the dummy equals one. The baseline values are adjusted accordingly. The endline values are used because the shock exposure during the intervention phase is of special interest to assess a heterogenous treatment effect. Equation (3) demonstrates the regression function including a shock in a difference-in-differences-in-differences (DiDiD) approach. The difference between the beneficiaries that experienced a severe shock and those who did not is δ_2. δ_1 becomes the treatment effect for households that did not experience a certain shock severely.

$$(3)\quad y_{it} = \beta_0 + \beta_1 * wave_t + \beta_2 * treat_i + \delta_1 * (wave_t * treat_i) + \boldsymbol{\beta_3 * shock_{it} + \beta_4 * (shock_{it} * wave_t) + \beta_5 * (shock_{it} * treat_i) + \delta_2 * (shock_{it} * treat_i * wave_t)} + u_{it}$$

Some shocks might be interrelated, e.g., crop loss through natural disasters could lead to more theft and conflict. If so, the model will not assess the different impacts properly[40]. One possible solution to include the effects of the different shocks is one aggregate indicator[41]. This is a plausible solution for objective data. The subjective data in this setting are likely to bias the comparability of the aggregate index because households perceive shock severity differently. The sum would exacerbate this bias assuming that the perception sensitivity is similar among different shocks. The separate evaluation reduces this bias.

[38] Through the differencing, the FE model works with half of the number of observations or less in case of attrition.
[39] In this work, the main analysis will make use of the data set excluding the attritted households. Therefore, the model is balanced and unbiased.
[40] The shock regressor will be correlated to the error.
[41] An index can be built out of the sum of experienced shocks, weighted by severity.

3.3 Descriptive

3.3.1 Data Overview

In this section, the data are presented in detail. Figure 6 gives an overview of the local dispersion of the observations. The left snapshot shows the interview locations of the baseline and endline observations. The clustering pattern, where the interviews were held in the first wave, are similar to those from the second wave with little exceptions. This underlines the reliability of the data set. On the right side, the locations of the control and the treatment households at the baseline are displayed. It becomes clear that in every cluster, both groups are represented and that the households live very close to one another. This is valuable for a precise counterfactual estimation because local exogenous influences that affect the development are probably not limited to just one group.

Figure 6: Distribution of Baseline and Endline Observations (left) & Distribution of Control and Treatment Observations at the Baseline (right)

Snapshots from own visualization with QGIS and OpenStreetMap

Table 3b in section 3.2.3 gives an overview of key socio-economic characteristics of the participants at the baseline before PSM[42]. At the baseline, the share of female respondents in the control group was significantly higher with 74% compared to 55% in the treatment group. This gap is remarkable. Approximately half of the household heads in both groups are female. The mean age of the respondents is in the early-mid-forty-year range, treatment respondents are on average forty-five years-old, while control households are on average forty years-old. The household heads are on average slightly younger with forty-four years in the control group and forty-six years in the treatment group.

Most of the respondents state that their household head is a farmer. This is more common in the treatment group with 95%, compared to the control group with 85%. This is clearly explained by the assignment to treatment of farming households. One third of the HHH of the whole sample are literate and around 80% of the household heads are married in the full sample, both variables are balanced between the two groups. The mean household size is around 8.5 in both groups. 10% of the control and 6.5% of the beneficiary households are IDPs or returned in the past five years from being displaced. The participating control household spends on average 50€[43] on food per week, around 10€ more than the treatment households. This is probably connected to the higher share of farming households within the treatment group. The mean agricultural land size is around .75ha for control households and .70ha for treatment households at the baseline. 95% of the households possesses a land size smaller than 2 ha[44]. 55% of the control households have land access to water at the baseline, while in the treatment group, only 45% have water access on their agricultural plot at the baseline. Around 20% of the beneficiary sample received humanitarian support apart from the present aid program during the intervention period, while in the control group only 6% received support in 2020. This is likely to be explained by the assignment of vulnerable households to treatment. During the intervention period 63% of the households were severely affected by the floods. This is balanced between the groups.

All in all, no alarming differences are detected, therefore the sample is roughly balanced. Even though, some of these socio-economic characteristics indicate significant differences between the groups, most of the differences are minor. This emphasizes that the sample is sufficiently powerful to detect small differences. The most severe disparities between the two groups are the respondents' gender, the heads of household being a farmer, land and water access and the reception of support. Differences in these variables possibly disturb the common trend. Therefore, the PSM approach increases the model's precision.

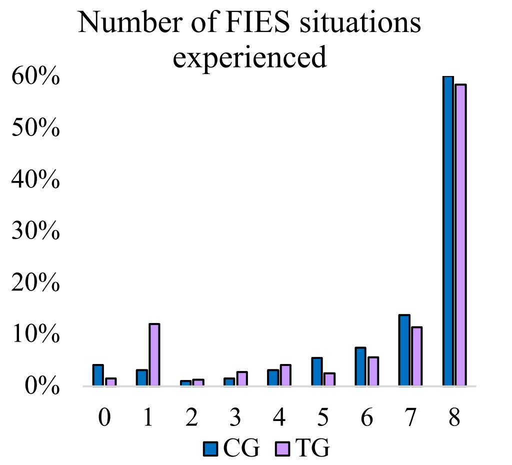
---

[42] In this section, the values are compared without PSM to elaborate the real trends and differences of the sample.
[43] Exchange rate from 16.12.2019 (0.0063) is used for wave 1 (Mataf, 2021).
[44] Information not displayed.

The food security outcomes at the baseline are alarming. Table 5 gives an overview of the baseline outcomes of the food security indicators before PSM. The first panel captures FIES. The left side presents the shares of respondents that experienced a certain number of

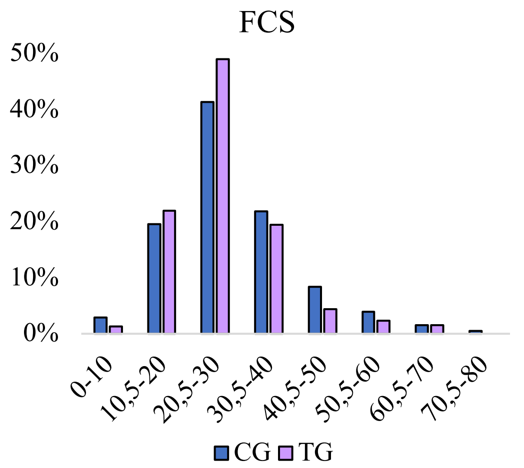
Table 5: Food Security Indicators at the Baseline

1. Food Insecurity Experience Scale



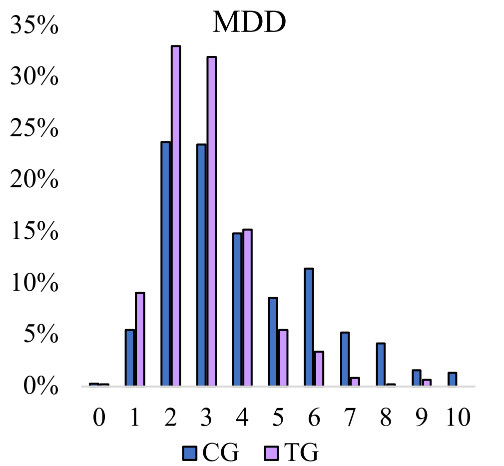
	CG	TG	p-value
% respondents experienced certain situations in 12 months			
worried	85.7	84.4	.618
healthy	84.6	79.3	.042
fewfood	88.5	81.0	.002
skipped	85.4	78.0	.005
ateless	85.9	80.6	.036
runout	83.6	78.0	.038
hungry	82.3	79.0	.233
whlday	76.8	74.3	.395
Food insecurity classification according to Rasch model			
% Moderate + severe	87.6	80.7	n.a.
% Severe	77.4	73.4	n.a.

2. Food Consumption Score



	CG	TG	p-value
Average No. of days per week of household consumptions per food group			
starches	5.8	6.3	<.001
Pulses	1.9	1.5	<.001
Vegetables	3.1	3.3	.170
Fruits	.5	.3	<.001
Meat, eggs, and fish	1.4	1.1	.014
Milk	.3	.3	.893
Sugar	1.4	1.1	.077
Oil	.9	.8	.170
Condiments	2.9	3.2	.112
FCS classification			
% Acceptable	22.4	18.8	.198
% Borderline	51.6	53.8	.521
% Poor	26.0	27.4	.650

3. Minimum Dietary Diversity



	CG	TG	p-value
% respondents consumed certain food group in one day			
grains, white roots and tubers, and plantains	95.6	98.5	.015
pulses	56.5	41.7	<.001
nuts and seeds	44.0	17.3	<.001
diary	14.3	9.3	.025
meat, poultry, and fish	33.6	22.9	<.001
eggs	6.5	3.2	.030
other vitamin A-rich plants	34.1	14.7	<.001
green, leafy vegetables	55.2	55.3	.908
other vegs	38.5	22.9	<.001
other fruits	12.8	8.0	.019
Higher micronutritional adequacy			
% with adequate level	32.3	10.8	<.001

Notes: p-values from t-tests with unequal variances (welch), CG=Control Group, TG=Treatment Group

FIES situations. All in all, the distribution is very similar when comparing both groups. The graph shows that around 60% of the respondents experienced all FIES situations in the past year before baseline, and only very few respondents did not experience any FIES situation in that time frame. Beneficiary respondents are more likely to experience only one FIES situation. On the right side, FIES is compared with more details. All situations are experienced by three-fourths or more of the sample. Running the whole day without food, the most severe item, is the least prevalent situation for both groups. The shares are slightly higher for control households. Accordingly, control group respondents are slightly more often classified as food insecure by the Rasch method, 77% of the control respondents are classified as severely food insecure and 88% as moderately or severely food insecure at the baseline. For the treatment group respondents, the shares are 73% and 81% respectively.

The second panel of Table 5 deals with FCS. The left side captures the indicator. The graph implies that, again, in both groups the households are similarly distributed on the FCS at the baseline. 41% of the control group and 49% of the treatment group have an FCS between 20.5 and 30, which is considered borderline to acceptable food consumption. On the right side, FCS is compared in detail. Starches are the mostly consumed food component, which is consumed nearly every day, followed by vegetables and condiments, which are consumed approximately every second day. The other groups are consumed less regularly. The FCS indicates that 20% of the sample households have an acceptable, 52% a borderline, and 27% a poor food consumption at the baseline. There are no significant differences in classification between the groups; however, the control group performs slightly better.

In the third panel of Table 5, the Minimum Dietary Diversity Score in described. The left side implies that the respondents from the control group eat more diversely at the baseline. The detailed illustration on the right side implies that nearly all respondents consumed grains, roots, tubers, or plantains one day before the baseline interview, and on around every second day pulses and green leafy vegetables. The control group respondents are more likely to consume nearly every category one day before the baseline survey. Exceptions are with grains, white roots and tubers, and plantains and green, leafy vegetables, the consumption of which is approximately similar in both groups. This is also reflected in the MDD classification: while 32% of the control group respondents meets the requirements for adequate micronutrition, this applies to merely 11% of the treatment group.

In conclusion, the food security show that differences between the two groups are present. The significance of the disparities stress again that it is reasonable to make use of PSM to counteract for selection bias. The indicators draw an ambiguous picture: While the treatment group performs better in FIES, it shows worse outcomes in FCS and MDD. The detailed comparison of the latter two indicators shows differences even though both assess the consumption of food groups. The differences are likely to be explained by the different food grouping approach, the different recall frame or by the fact that FCS is assessed on household level, while MDD is captured on the individual level. This implies intra-household differences in dietary diversity. The different conclusions imply the multidimensionality of food security and the importance to assess it from diverse perspectives.

In the next step, the sample comparison will shed light on adverse shock experiences. The development from the second wave will be included, to monitor the common trend. Table 6 illustrates the share of households that state to have experienced selected shocks, severity is not regarded in this part. Table 6 emphasizes that all seven shocks are prevalent in both waves and in both groups. The most prevalent shocks at the baseline which are experienced by more than 75% of the households are high agricultural costs, crop pests, and theft (only for the beneficiaries), while floods are the most prevalent shock impact in the second wave, which are experienced by over 80% of the sample. In the first wave, every treatment household experienced at least one of the shocks in the past twelve months, while this is true for 98% of the comparison group. In the second wave, 7% of the control group and 3% of the treatment group did not experience any shock.

Table 6 indicates that at the baseline a significantly higher share of treatment households is targeted by floods, conflict and violence, theft, and high crop pests compared to the control group. In the second wave, significantly more treatment households are affected by floods, high agricultural costs, and theft. It becomes clear that the certain adverse shocks are more prevalent in the treatment group, which is probably explained by different levels of exposure by neighborhood, by occupational disparities, or by different levels of resilience[45]. Overall, the trends in both groups are mostly comparable.

For this analysis, it is vital that agricultural key figures would have developed similarly in the absence of treatment. Agricultural practices should be comparable at the baseline. However, different developments in agricultural success might be attributed to the program. The groups possess a quite similar land size on average at baseline (see Table 3b). Table 7 stresses that in both groups, most of the households cultivate land, however, this share declines drastically from 86% to 57% for the control group within one year. Farming households possess on average .1 ha of arable land per household member at the baseline. This ratio also declines substantially for the control group, while it remains approximately stable for the treatment group. Considering that the treatment included a land expansion, this indicates that the arable land size declined over time, probably because of the floods. The control group shows higher values in harvest at the baseline, which are significantly

[45] As this score is a subjective measure, more resilient household are less likely perceive to have experienced certain shocks if they are comparatively less affected.

different from the treatment group in cereals and fruits. Over time, all categories decline for the control group. The treatment group shows again more stable indicators.

Reviewing the data overview, the severe situation of the survey participants is stressed. Many of the participating households and respondents are in alarming food security

Table 6: Adverse Shock Experience

Percentage of Sample that experienced selected adverse shocks in the past 12 months

	wave 1			wave 2		
	CG	TG	p-value	CG	TG	p-value
Drought	63.8	65.4	.620	26.3	26.6	.931
Floods	43.5	70.4	<.001	83.1	89.0	.014
High agr. costs	84.1	86.0	.455	51.6	57.7	.076
High crops pests	80.2	84.9	.076	57.0	61.1	.228
Illness/accident	48.4	43.4	.145	33.6	34.1	.871
Conflict/violence	22.4	31.5	.003	24.5	23.8	.807
Theft	60.4	76.9	<.001	37.2	48.2	.001
No shock	1.8	0	.008	6.5	3.2	.030

Notes: p-values from t-tests with unequal variances (welch), CG=Control Group, TG=Treatment Group

Table 7: Key Figures of Agricultural Production

	wave 1			wave 2		
	CG	TG	p-value	CG	TG	p-value
% HH that cultivated land	85.7	88.0	.343	56.8	85.5	<.001
Arable land/HH member	.11	.09	.053	.05	.08	<.001
Mean harvest in kg						
Cereals	72.8	51.0	<.001	16.3	47.0	<.001
Other main crops	71.8	63.3	.188	18.4	44.6	<.001
Vegetables	39.3	34.2	.236	16.9	45.0	<.001
Fruits	21.4	12.4	.005	2.7	15.1	<.001

Notes: p-values from t-tests with unequal variances (welch), CG=Control Group, TG=Treatment Group, Values are set 0 for households that do not operate in agriculture.

conditions, are frequently exposed to adverse shocks, and the agricultural key figures declined notably for the control group. It becomes obvious that the two household groups are mainly comparable. There are some disparities in socio-economic features, in outcome variables, and in shock exposure between the two groups, which are addressed by the model.

3.3.2 Data Quality

In the questionnaire, most of the questions are mandatory. The only non-mandatory section is food security because the questions are sensible. The idea is to prevent the respondents from feeling uncomfortable giving details on topics they are ashamed of. This computation is used to prevent the respondents from giving false answers if they are unwilling to answer. Each of the food security questions has less than 1% of unanswered responses. The non-responses are computed as 0, as is most cases only particular questions of the food security section are not answered[46]. Therefore, this could also be an accidental non-response. As the non-response is very low the missing data are not regarded as a possible cause of bias.

Another possible pitfall is systematic false response out of numerous reasons, e.g., if a respondent answers a complete section with *"no"* if indeed *"yes"* is the true answer[47]. The monitoring of this behavior is not possible. The bias is less threatening if there are no systematic differences between the groups and the systematic false response happened on a small scale. Measurement errors are restricted due to the survey tool programming. However, small mistakes cannot be avoided and might lead to an attenuation bias but if this is random, the expected harm is minor.

Table 8: Data Quality Evaluation

Percentage of panel households that gave same responses in both waves	
HHH gender*	85%
HHH literacy*	59%
HHH marital status*	55%
HHH farmer*	88%
IDP/returnee HH	85%
HH size (±2)	44%

Notes: *Only included if the HHH did not change. HH=Household, HHH=Household head

A high within-household variety in time-invariant characteristics indicates a lack of reliability of the data. To monitor the quality, Table 8 summarizes the share of households, who gave the same responses in basic characteristics that are assumed to be time-invariant between the two waves. The assessment of the household head characteristics is only included if the household did not change between the two waves[48]. The table indicates that 85% identify their household head with the same genders in both waves, 59% classify their household head's literacy equally, 55% give the same marital status of the household head, and 88% give the same response on whether the head of the household is a farmer in both periods. 85% of the respondents state the same residency status of the household, and 44%

[46] The unanswered sections are not counted as missing because mostly only one or two responses are missing per section. This would lead to a high observation loss.
[47] The atmosphere of 97% of the interviews is rated as *"good"* or *"very good"* by the enumerators. That respondents and enumerators felt comfortable with each other, is assumed to reduce the risk of systematic false response.
[48] 88% of the respondents stated the same name as their household head in both waves.

respond with a similar number of household members (±2). These numbers appear to be in high flux and indicate that the responses should always be questioned for plausibility. It is important to be aware of possible data reliability issues as the survey was self-reported. Apart from inevitable measurement errors, some variation is plausible as these variables might indeed change over time or they are interpretable differently, e.g., the definition of household members living in the household if relatives live next door. Especially, in a setting with high mobility, the household size might change frequently, wherefore the interpretation of the residency status could also differ. Another obstacle is the high rate of illiteracy, which is likely to correlate with erroneous number statements. In the end, the responses must be assumed to be correct.

4 Main Findings

4.1 The Impact of Small-Scale Agricultural Support on Food Security

4.1.1 Food Insecurity Experience Scale (FIES)

Figure 7a: Distribution of Number of experienced FIES Situations Control Group

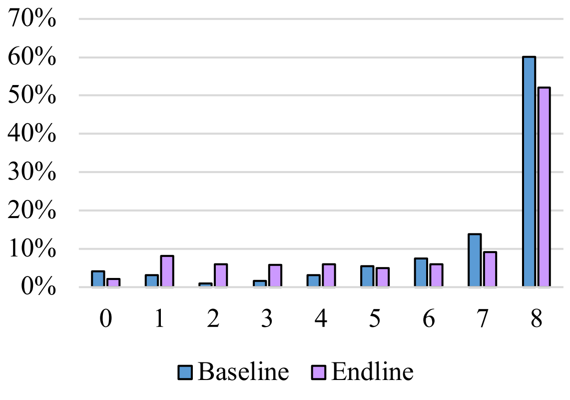
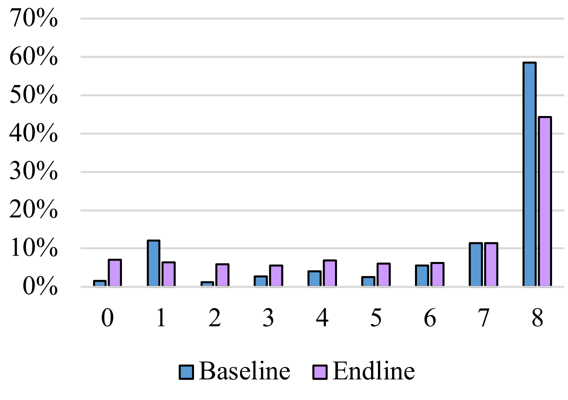


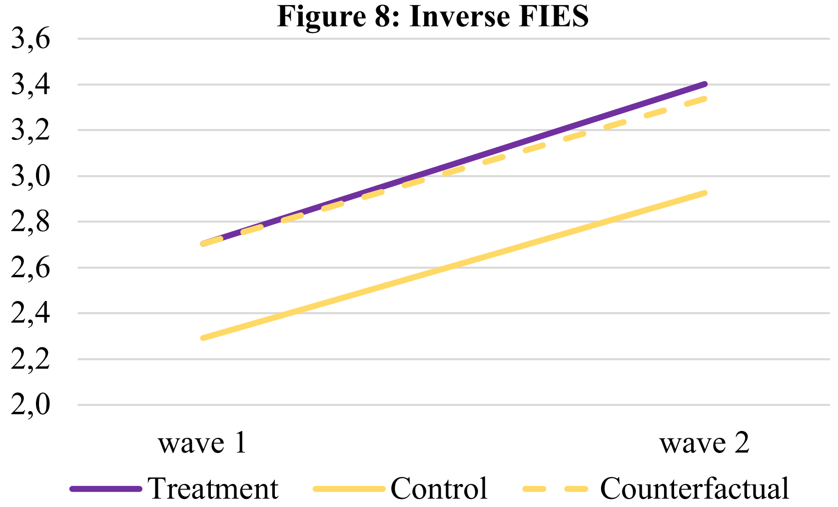
Figure 7b: Distribution of Number of experienced FIES Situations Treatment Group



The main analysis begins with the Food Insecurity Experience Scale. Figures 7a and 7b display the distribution of respondents that experienced a certain sum of FIES situations by wave and by control and treatment group, respectively. The figures indicate that the shares of respondents that experienced all eight FIES items in the past twelve months decreased from baseline to endline by 8% points and by 15% points, for the control and treatment group, respectively. Still, at the endline, 52% and 44% of respondents respectively experienced all items. The share of beneficiary households that did not experience any FIES situation increased from 2% to 7%. These figures indicate a positive development, even though the overall situation remains severe. The beneficiary group shows stronger improvements.

In this thesis, the ATT is the main indicator of the program impact. To evaluate this, I compare the outcome variable development of the control and the treatment group. DiD plots illustrate the average changes between the two waves for both groups and the counterfactual of the treatment group. The distance between the counterfactual and the treatment group outcomes in the second period is the average treatment effect of the treated.

Figure 8: Inverse FIES



Notes: Control Group is PSM adjusted.

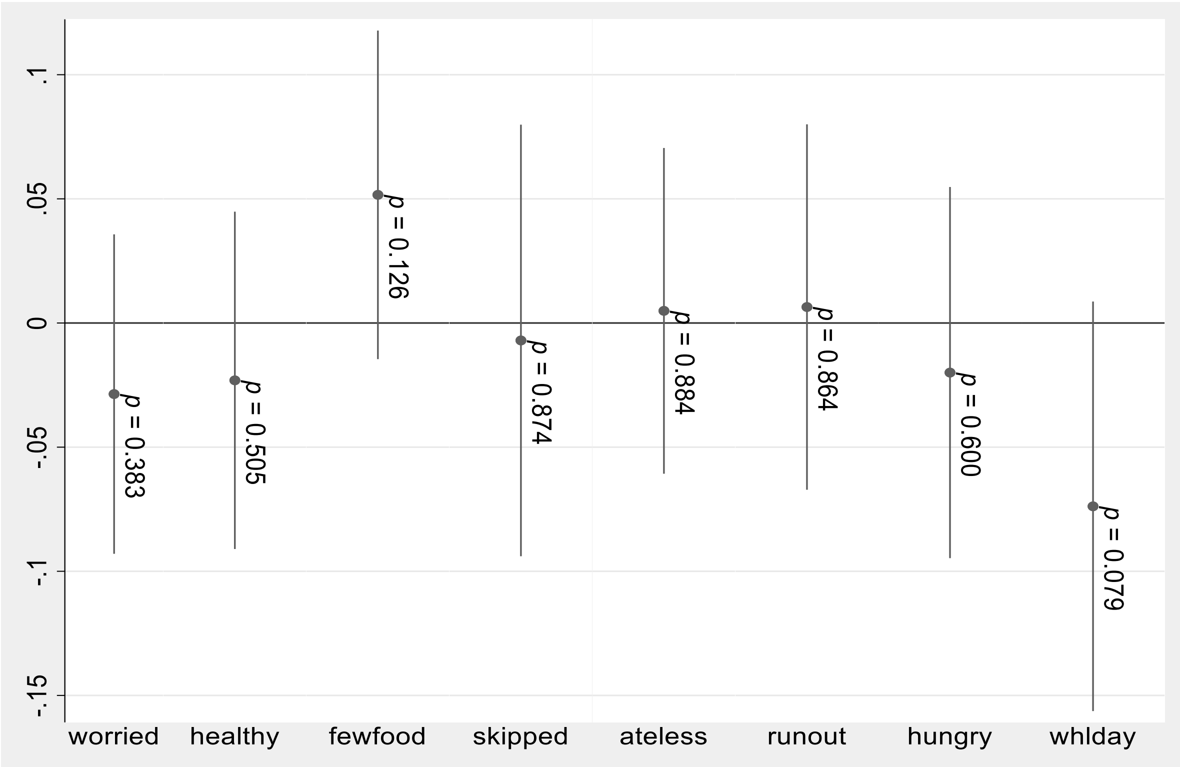
The inverse FIES spans from 1-9, where 9 means that none of the FIES situations was experienced in the past twelve months, and 1 means that all eight situations were experienced. Figure 8 displays the DiD approach with kernel-based PSM for this scale. At the baseline, the control group reaches an average inverse FIES score of 2.3, while the treatment group reaches a slightly higher mean score of 2.7. This gap increases from .4 to .5 points in the endline. Both groups reach slightly better average outcomes of 2.9 and 3.4, respectively. Therefore, the food security situation overall improved on average for both groups by .6 and .7 points, respectively. This means that on average 60-70% of the respondents experienced one FIES situation less in the past twelve months before the second interview compared to the past twelve months before the first interview. There is a very small ATT of .1 visible. Table 9 displays the findings in detail. The ATT is .06 with three times higher SE. Therefore, the ATT is as good as zero. Consequently, I find that no direct impact of the intervention on improvements in food security is measured by FIES.

Table 9: Inverse FIES & Treatment

	Baseline			Endline			ATT
	CG	TG	Diff	CG	TG	Diff	
Inverse FIES	2.242	2.654	.413** (.187)	2.876	3.352	.476** (.208)	.064 (.205)

Notes: N=1,694, CG=Control Group, TG=Treatment Group, Diff=Group Difference, Pooled OLS regressions, Kernel-based PSM, SE in parentheses and clustered on household level, * p < .1, ** p < .05, *** p < .01

Figure 9: The ATT on FIES Components



Notes: Average treatment effects with 95%-Confidence intervals (CI) on FIES components with linear probability model. SE clustered on household level. Control group weighted by kernel-based PSM. The dots indicate the ATT, the lines display the CIs. If the CI does not cross the line at zero, the effect can be attributed to be significant at the 5% level. The ATT displays changes in the likelihood to experience a certain FIES situation.

The DiD analysis implied an absence of a significant ATT on FIES. However, it is possible that treatment had different influence directions on certain components of the scale so that the overall effect diminished. To analyze these different nuances of the FIES score, I run a linear probability model with the DiD and PSM framework on the eight FIES questions separately. Figure 9 allows to compare the ATT on each of the outcome components including the 95%-confidence intervals (CI). The figure implies that treatment decreased the likelihood to experience a whole day without food by 7% points (p=.08), whereas it can be concluded that treatment decreased the odds to experience the most serious food insecurity situation. Treatment tends to increase the probability to only have access to few food components by 5% points. This is likely to be related to the awareness of adequate nutrition created by the intervention. The effect on other FIES components is very close to zero and insignificant, so that even if I were to have proved significant effects with a larger sample, the effects would be barely notable in society.

4.1.2 Food Consumption Score (FCS)

Figure 10a: FCS Classification Control Group

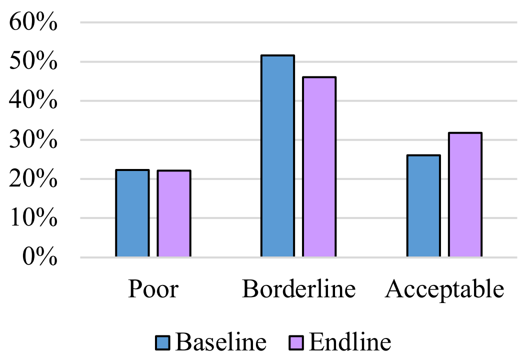
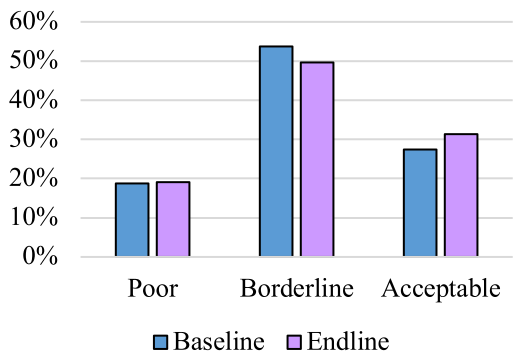


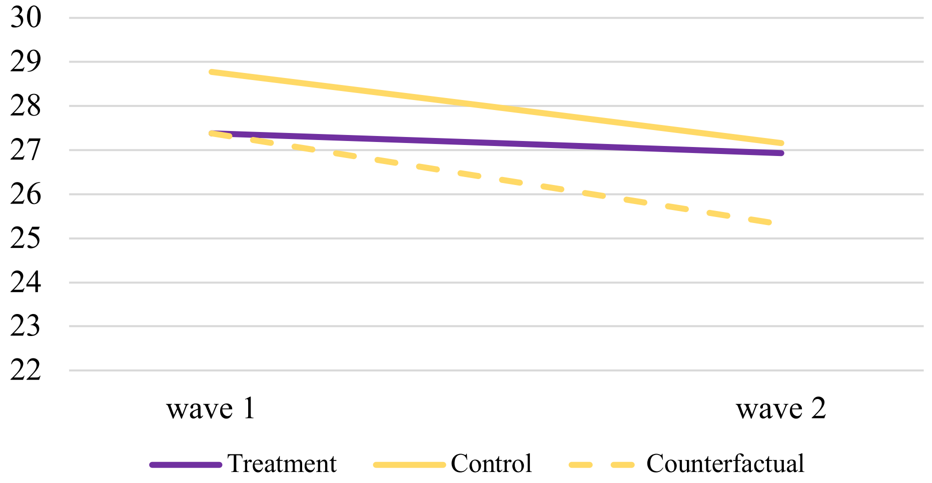
Figure 10b: FCS Classification Treatment Group



Next, the analysis will concentrate on the treatment effect on dietary diversity. The Food Consumption Score ranges from 0 to 112 and captures the household consumption of different food groups weighted by its nutritional importance. Figures 10a and 10b display the distribution of the households in the FCS classifications by wave and by control and treatment group. The figures indicate that around 20% of the participating households is classified by poor food consumption. This share remained constant in both waves. Around half of the households are characterized by borderline food consumption. This share decreases from baseline to endline by 6% points and 4% points in the treatment and control group, respectively. Accordingly, the share of households with acceptable food consumption increased from 26% to 32% in the control group and from 27% to 31% in the treatment group. The two groups show very similar outcomes and developments. Like the FIES, the FCS indicates a severe but nonetheless improving trend.

Again, Figure 11 illustrates the DiD approach. The mean FCS decreased slightly for the control group from 29 to 27 points, while it remained on the same level of around twenty-seven points for the treatment group. This implies a positive ATT but regarding the range of the scale from 0-112, the differences are very small. Table 10 indicates a small positive and insignificant ATT in FCS in the first model of one point. This value is very low. Therefore, it can be concluded that treatment did not influence the food consumption score

Figure 11: FCS



Notes: Control Group is PSM adjusted.

substantially.

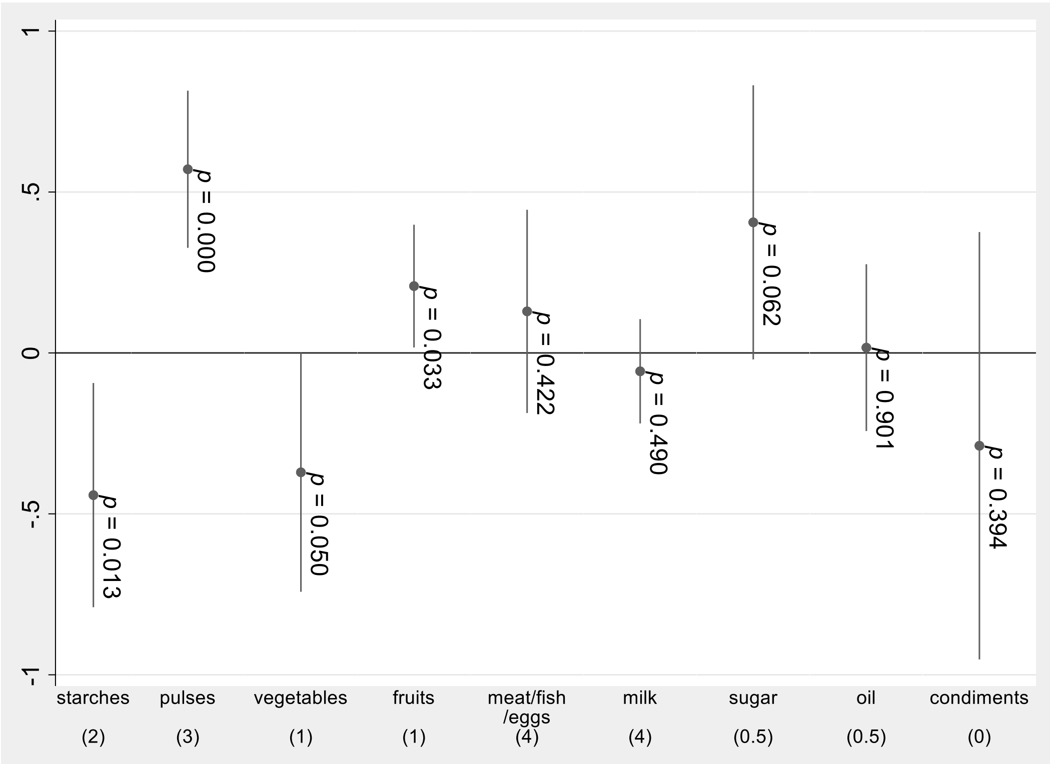
Table 10: FCS & Treatment

	Baseline			Endline			ATT
	CG	TG	Diff	CG	TG	Diff	
FCS	28.774	27.380	-1.394 (.913)	27.159	26.931	-.228 (.913)	1.165 (1.184)

Notes: N=1,694, CG=Control Group, TG=Treatment Group, Diff=Group Difference, Pooled OLS regressions, Kernel-based PSM, SE in parentheses and clustered on household level, * p < .1, ** p < .05, *** p < .01

There might be dynamics in the consumption composition, which did not affect the score substantially because of the weighting or because of counteracting components. Figure 12 illustrates separate DiD regressions with the PSM framework on the components of the FCS. The range is +/-7, as the score accounts for days with intake of the certain categories in one week. In this part, I ignore the nutritional importance weights in parentheses and concentrate only on the dietary composition. A significant increase of consumption in pulses (p<.01), fruits (p=.03), and sugar (p=.06) through treatment is observable while the intake of starches (p=.01) and vegetables (p=.05) decreased significantly. For meat, fish, eggs, milk, oil, condiments, no unambiguous effect is detectable. The largest impact is assessed for the increase in the consumption of pulses: Through the program, more than every second household consumed pulses on one additional day in the week before the second survey.

Figure 12: The ATT on FCS Components



Notes: Average treatment effects with 95%-Confidence intervals (CI) on FCS components with OLS regression. SE clustered on household level. Control group weighted by kernel-based PSM. Relative nutritional importance in parenthesis. The dots indicate the ATT, the lines display the CIs. If the CI does not cross the line at zero, the effect can be attributed to be significant at the 5% level. The ATT displays changes in the average frequency of consumed food groups. The range is +/-7.

4.1.3 Minimum Dietary Diversity (MDD)

Next, I focus on Minimum Dietary Diversity. The MDD score ranges from 0 to 10 assessing the number of different food groups consumed by the respondents on the previous day. An intake of five or more food groups indicates a higher micro-nutritional adequacy. Figures 13a and 13b display the classification development. The share of the respondents who have a MDD ≥5 in the control group decreased from 32% to 18% within one year, while this share increased within the treatment group from 11% to 21%. The developments clearly differ between the two groups, indicating a positive ATT.

Figure 13a: Share of Respondents with higher micronutritive Adequacy Control Group

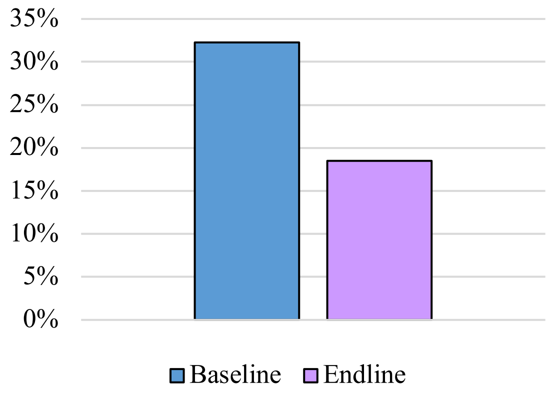


Figure 12b: Share of Respondents with higher micronutritive Adequacy Treatment Group

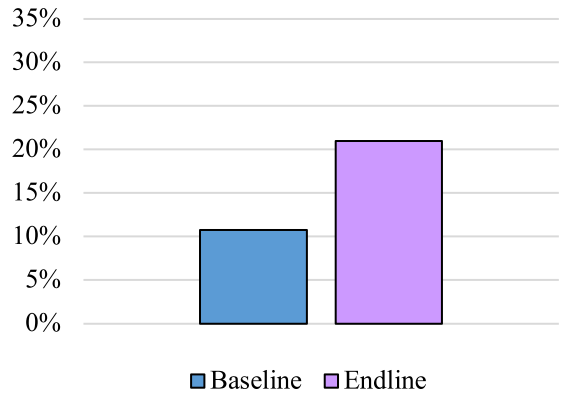
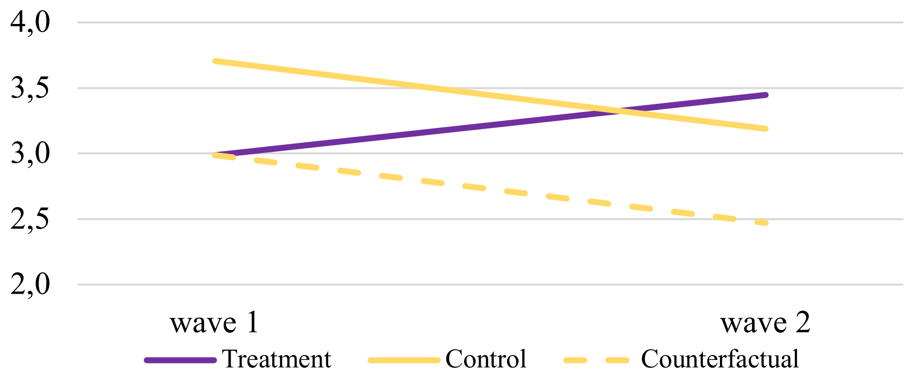


Figure 14: MDD



Notes: Control Group is PSM adjusted.

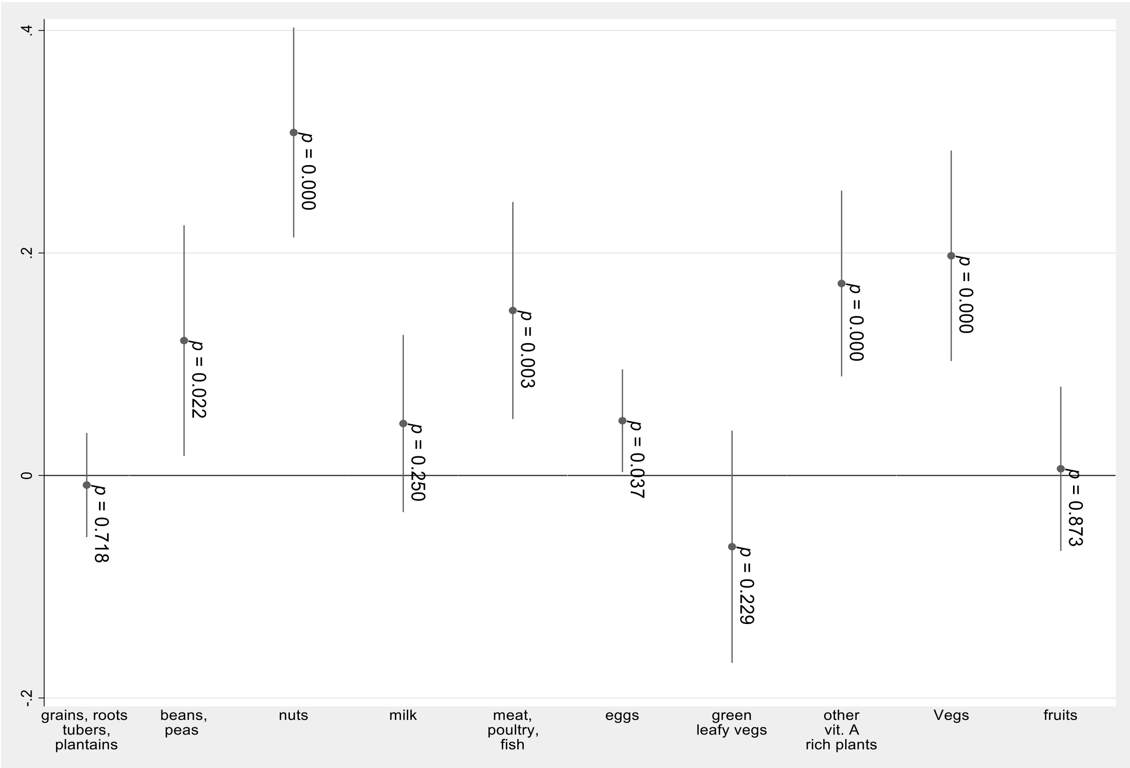
Table 11: MDD & Treatment

	Baseline			Endline			ATT
	CG	TG	Diff	CG	TG	Diff	
MDD	3.656	2.937	-.719***	3.139	3.397	.258**	.977***
			(.131)			(.129)	(.168)

Notes: N=1,694, CG=Control Group, TG=Treatment Group, Diff=Group Difference, Pooled OLS regressions, Kernel-based PSM, SE in parentheses and clustered on household level, * p < .1, ** p < .05, *** p < .01

The DiD plot in Figure 14 assesses the MDD score. It emphasizes a negative time trend in MDD from 3.6 to 3.1 consumed food groups on average on the pre-survey day for the control group, while for the treatment group, the score increased from 2.9 to 3.4. This leads to a clear gap between the outcome for the treatment group and its counterfactual at ~1 point on the MDD scale. This means that the beneficiary respondents consumed on average one additional food group on the previous day through treatment. Table 11 confirms the conclusions from Figure 14. The estimate indicates an increase in MDD on average by one point (p<.01). Considering the SE, the impact is very stable.

Figure 15: The ATT on MDD Components



Notes: Average treatment effects with 95%-Confidence intervals (CI) on MDD components with Linear Probability Model. SE clustered on household level. Control group weighted by kernel-based PSM. The dots indicate the ATT, the lines display the CIs. If the CI does not cross the line at zero, the effect can be attributed to be significant at the 5% level. The ATT displays changes in the likelihood to consume a certain food group.

Figure 15 again illustrates the breakdown analysis of the MDD components. The ATT on the likelihood for the respondents to consume certain food groups is assessed with a Linear Probability Model. The figure implies a tendency towards a decrease in the probability of green, leafy vegetables consumption on the previous day. Moreover, the figure emphasizes significant increases in consumption of beans and peas (p=.02), nuts (p<.01), meat, poultry, and fish (p<.01), eggs (p=.04), other vitamin-A-rich fruits and vegetables (p<.01), and other vegetables (excluding green, leafy vegetables) (p<.01) through treatment. Nuts and the latter two groups especially show a clear increase of consumption by 17-31% points. The probability of consuming milk tends to increase through treatment as well. Interestingly, the consumption of grains, roots, tubers, and plantains, and of fruits does not change notably through treatment contradicting the conclusions from FCS, where the consumption of starches decreased, and the intake of fruits increased. This is probably explained by the different grouping. MDD assesses the groups more meticulously. Therefore, the consumption of vegetables appears to decrease only in green leafy vegetables but increase in other types. MDD assesses the effect on the individual level while FCS is addressed to the household level. Therefore, the ATT probably differs on the intra-household level. All in all, MDD supports mainly the conclusions from FCS of a treatment effect on the dietary composition.

4.2 Treatment Components

The previous section showed that the combined intervention, consisting of different agricultural support tools and a cash transfer, did not lead to substantial impacts on FIES and FCS, while MDD increased substantially. The monetary support was provided to severely flood-affected beneficiary households in August 2020. This is likely to have had a substantial impact on food security. For the project evaluation, it is essential to distinguish whether the impact on dietary composition is caused by the agricultural measures or by the one-time payment. The transfer can be spent on food or other goods. Assuming that the transfer is not invested in agriculture, the only pathway to affect food security in the short run is through purchasing more food. On the contrary, the in-kind support combined with the capacity building program, is assumed to stimulate food security indirectly through agricultural productivity.

4.2.1 Cash Transfers versus SAI

Table 12 shows DiD regressions only including the beneficiary households. The treatment *"cash"* now distinguishes between beneficiary households that received a cash transfer and those who did not. The DiD regressor, therefore, assesses the ATT of the cash transfer, differentiating out the impact of the agricultural components. For more accurate comparability, the group that did not receive the transfer is again weighted with kernel-based PSM[49]. It is important to regard that the cash transfer also indicates that the household is affected by the floods in 2020, wherefore the comparability of the subgroups is lacking in terms of isolating the clean cash impact. The sign of the DiD estimator gives information on whether the impact of the floods surpasses the cash impact[50]. In this part, logarithmic average weekly food spending is included as an outcome to learn about the impact of the transfer on spending behavior. This part of the analysis must be regarded with caution be-

[49] The pre-and-post PSM group balance is displayed in Table A4 in the appendix. The t-tests overall do not indicate substantial differences. This is partially attributed to the small control group. Therefore, the PSM approach from 4.1 is applied to ensure the maximum potential of comparability. Food spending and flood experience are excluded from PSM to avoid endogeneity. Land size is added because it differs substantially between the groups. Five observations of the group that did not receive the transfer drop out because they do not match sufficiently. Common support is given (see Figure A3 in the appendix).
[50] Under the assumption that cash has a positive impact and floods have a negative impact on food security.

cause the treatment group that did not receive the transfer only includes fifty-eight observations.

Table 12: The Impact of the Cash Transfer (Subsample: Beneficiary Households)

	Baseline			Endline			
	No Cash	Cash	Diff	No Cash	Cash	Diff	ATT
inverse FIES	1.378	2.835	1.457***	2.505	3.514	-1.008***	-.428
			(.213)			(.381)	(.342)
FCS	27.244	27.512	.268	26.67	27.044	-.374	-.107
			(1.429)			(1.231)	(1.730)
MDD	3.102	2.914	-.189	3.247	3.462	-.215	.404
			(.275)			(.224)	(.318)
ln (weekly food spending)	3.270	3.334	-.064	3.424	3.344	-.079	-.143
			(.156)			(.168)	(.209)

Notes: N=916, No cash=treatment households that did not receive the transfer, cash=treatment households that received the transfer. Pooled OLS regressions, Kernel-based PSM, SE in parentheses and clustered on household level, * $p < .1$, ** $p < .05$, *** $p < .01$

Table 12 indicates that none of the outcomes show a significant ATT, which means that there are no clear differences in the treatment effects between the subgroups. The signs of the DiD estimators are negative for all outcomes but for MDD. The cash transfer combined to the severe flood experience tends to have a small positive but insignificant impact on MDD. In contrast, there is a slight but insignificant tendency of a negative ATT on inverse FIES, this indicates that the negative impact of the flood experience seems to have exceeded the positive cash transfer-shock in this indicator. These controverse indications could be explained by the time frame of FIES, which covers twelve months and, therefore, covers at least seven months without the transfer, while MDD recalls the previous day. Additionally, the floods started before the cash distribution. Still, the ATTs are very low. The DiD estimator for FCS is small combined to high SE, therefore, the effect is as good as zero. The regression on food spending implies that households, who received the cash transfer spend roughly the same on food like those who did not receive the transfer in both waves. Therefore, the impact of cash on food expenditure is as good as zero. This suggests that the households spent the received money on other things, likely on the reconstruction of the flood-destroyed assets. It is worth noting, as the distribution took place three to four months before the interview assessment, that the money might be already spent on food and disregarded in the survey. Taking the positive tendencies in MDD into account, this seems imply that the transfer shifted the expenditure towards a more diverse diet without increasing the expenditure. This leads to the assumption that the transfer was indeed invested in agriculture[51].

Table 13: The Impact of the Combined Treatment (Subsample: Control Group and Treatment Households that received the Transfer)

	Baseline			Endline			
	CG	Restr. TG	Diff	CG	Restr. TG	Diff	ATT
inverse FIES	2.180	2.835	.655***	2.828	3.514	.685***	.030
			(.188)			(.212)	(.201)
FCS	28.609	27.512	-1.096	26.770	27.044	-.274	1.371
			(.902)			(.908)	(1.126)
MDD	3.702	2.914	-.788***	3.161	3.462	.300**	1.089***
			(.136)			(.137)	(.181)
ln (weekly food spending)	3.417	3.334	-.083	3.401	3.344	-.057	-.027
			(.089)			(.099)	(.122)

Notes: N=1,570, CG=Control Group, restr. TG=Treatment Group only including households that received the cash transfer, Diff=Group Difference, Pooled OLS regressions, Kernel-based PSM, SE in parentheses and clustered on household level, * p < .1, ** p < .05, *** p < .01

Table 14: The Impact of the Agricultural Component (Subsample Control Group and Treatment Households that did not receive the Transfer)

	Baseline			Endline			
	CG	Restr. TG	Diff	CG	Restr. TG	Diff	ATT
inverse FIES	2.284	1.397	-.888***	2.928	2.224	-.704**	.183
			(.208)			(.298)	(.290)
FCS	28.288	26.457	-1.931	27.140	26.138	-1.002	.829
			(1.376)			(1.444)	(1.766)
MDD	3.720	3.103	-.616**	3.305	2.948	-.356	.260
			(.248)			(.226)	(.319)
ln (weekly food spending)	3.385	3.192	-.193	3.338	3.234	-.104	.089
			(.144)			(.160)	(.192)

Notes: N=882, CG=Control Group, restr. TG=Treatment Group only including households that received the cash transfer, Diff=Group Difference, Pooled OLS regressions, Kernel-based PSM, SE in parentheses and clustered on household level, * p < .1, ** p < .05, *** p < .01

To explore if the combined treatment effect differs from the ATT of the agricultural intervention, the two beneficiary subgroups are matched to the control group with kernel-based PSM[52]. Table 13 starts with the part of the treatment group that received the transfer[53]. The ATT are very similar for inverse FIES, FCS, and MDD to the main results, so that the conclusions do not differ. That these ATT do not differ substantially from the main analysis is

[51] This assumption cannot be proven as the data do not allow to analyze the transfer spending further.
[52] The subgroup pre-and-post PSM balance is displayed in Tables A5 and A6 in the appendix. The blue shaded covariates are assigned to PSM. Common support is given (see Figures A4 and A5 in the appendix).
[53] Four observations from the control group drop out as they did not match in the PSM approach.

explained by the high share of treatment households that received the transfer. The last row emphasizes that the impact of the combined treatment, including the transfer, is as good as zero on food expenditure.

Table 14 reflects the same analysis, now only including the beneficiary households that did not receive the transfer[54]. Hence, only the impact of the in-kind and capacity building support is assessed. Here, the ATT are very similar to the previous effect for inverse FIES and FCS. This stresses that the agricultural intervention, apart from the money, did not influence these two outcomes substantially in the short run. Also, the agricultural intervention effect on food spending approaches close to zero. The positive impact of the combined program on MDD diminished drastically supporting the assumption that the intervention including the cash rather improved MDD instead of the agricultural intervention. This supports the hypothesis of the emergency aid being spent on agricultural investments.

In conclusion, for inverse FIES and FCS, the subgroup analysis did not detect major differences between those beneficiaries that received the additional cash transfer and those who received only agricultural support. Interestingly, the positive ATT on MDD appears to be attributed to the combined intervention, as it diminishes if only including the restricted treatment group that did not receive the additional cash transfer. This emphasizes that the positive cash impact probably exceeds the negative flood impact on MDD even though food spending was not influenced significantly. This also implies that the aid is probably invested more efficiently in providing transfers to the households directly. FCS indicates that the household intake of sugar increased which is rather connectable to the cash transfer than to the agricultural support (see Figure 12). That cash receiving households are not substantially worse off than others, indicates that the flood damage is muted by the transfer, which is a desirable outcome. Still, it must be considered that shock exposure is generally high in the whole sample, whereas the flood experience is but one of many severe adverse shocks that sampled households experience. The very small sample size of the treatment households that did not receive the transfer is too low to make robust conclusions.

4.2.2 Treatment, Production, and Food Security

The agricultural treatment is supposed to have led to an increase in agricultural production if it was successful. Table 15 shows DiD regressions on agricultural production of the four crop groups in kg, again with the PSM extension. A positive and significant ATT on production is observable for all groups. Treatment increased the harvest in kg per household on average by 50kg for cereals, by 19-28kg each for other main crops, vegetables, and fruits ($p<0.01$). If investments in agriculture after the cash transfer do not impact food security, this indicates that the agricultural part of the intervention had a positive impact on

[54] One observation from the control group drops out as it did not match in the PSM approach.

agricultural productivity. However, looking at the absolute harvest in kg, this table does not stress that farming became significantly more productive: the harvest of cereals and other main crops decreased by 8% and 29%, respectively for the treatment group and for vegetables and fruits, the harvest increased clearly by 32% and 25%, respectively. The control group's harvest declined dramatically. This suggests that the treatment group was more stable in a general trend of decreasing productivity.

Table 15: The Treatment Effect on Agricultural Production in kg

	Baseline			Endline			ATT
	CG	TG	Diff	CG	TG	Diff	
Cereals	77.546	50.959	-26.587***	23.082	46.982	23.898***	50.485***
			(8.926)			(6.961)	(9.228)
Other main crops	71.519	63.317	-8.201	25.245	44.639	19.394***	27.596***
			(8.127)			(6.369)	9.989
Vege-tables	40.804	34.220	-6.583	24.638	44.998	20.360***	26.944***
			(5.504)			(5.372)	(6.091)
Fruits	21.868	12.369	-9.499**	5.507	15.065	9.558***	19.057***
			(4.255)			(3.522)	(4.539)

Notes: N=1694, CG=Control Group, TG=Treatment Group, Diff=Group Difference, Pooled OLS regressions, Kernel-based PSM, SE in parentheses and clustered on household level, * p < .1, ** p < .05, *** p < .01

Table 16: Crop Production on Food Security Outcomes (Harvest in 100 kg)

wave	Cereals	Other main Crops	Vegetables	Fruits
Inverse FIES	.298***	.117**	.187*	-.046
	(.083)	(.076)	(.100)	(.122)
FCS	1.448***	1.327***	1.769***	2.871**
	(.458)	(.483)	(.679)	(1.110)
MDD	.181***	.153**	.212**	.524***
	(.069)	(.069)	(.116)	(.112)

Notes: N=1694, CG=Control Group, TG=Treatment Group, Diff=Group Difference, SE in parentheses and clustered on household level, * p < .1, ** p < .05, *** p < .01
Every cell reflects a different regression. The crop types are regressed on inverse FIES, FCS, and MDD with a Fixed Effects regression including the variable *"wave"*, which is not displayed here because it is trivial in this context and impedes a coherent visualization.

To assure that the increase in agricultural productivity can be associated with food security, the harvest of the four crop groups in 100 kg is regressed on the main outcomes with fixed effects regressions in Table 16[55]. The coefficient values indicate positive relationships of all food groups on all outcomes, except for fruits on inverse FIES, where as good as no relationship is assessed. Especially, an increase of cereals harvest is associated with significant improvements of all indicators (p<.01). MDD and FCS increase significantly with an increase in harvest in all categories. Higher scores in inverse FIES are also associated with

[55] Because the treatment and control group are not considered in this model, DiD and PSM are not applied.

a larger harvest in other main crops (p<.05) and vegetables (p<.1). All in all, looking at the magnitude of the coefficients, it becomes clear that even if the increase of agricultural production has a causal impact on food security, the impact is still low. Assuming causality, an increase of cereal harvest by 500 kg, increases the inverse FIES by 1.5, the FCS by 7.2, and the MDD by .9 points. This is a very small improvement for a harvest, which is beyond the cleaned maximum harvest in cereals of 300 kg per year in wave 2[56]. This underlines again, that the positive impact in MDD cannot be exclusively attributed to the SAI but rather to the combined treatment.

These conclusions help to explain why there is no ATT on inverse FIES and FCS assessable, even though the program enhanced agricultural production and agricultural production is clearly associated with better food security outcomes. The increases in agricultural production are not strong enough to substantially improve these outcomes.

4.3 Shocks, Treatment, and Food Security

Chapter 2.1.2 emphasizes that in several studies adverse shocks are assumed to worsen food security and agricultural production. Chapter 3.3.1 implies that the shares of the household groups that were exposed to adverse shocks differ partially. For a robust interpretation of the treatment effect, it is vital to know if such events induced a heterogenous treatment effect. This question is addressed with a DiDiD approach[57,58]. Table A7 in the appendix explores the heterogenous treatment impact for the seven selected shocks separately. The interaction term *shock*wave*treat* displays differences between treatment households that experienced a certain shock severely and those who did not, this is the DiDiD estimator. The DiDiD estimators for each of the shocks are summarizes in table 17.

Generally, nearly all DiDiD estimators indicate insignificant differences, which emphasizes that in this setting, there are seldom clear heterogenous treatment effects induced by severe adverse shock experience in the past twelve months. For MDD, the idiosyncratic shock *"illness/accident"*, which assesses if the household head suffered from a severe physical impairment, indicates significant differences between the severely- and non-severely affected beneficiary households (p<.01). Treatment households that experience this shock during the intervention period severely, loose on average 1.1 MDD points, compared to non-severely affected beneficiary households. This absorbs the positive ATT on MDD, so that the ATT of the intervention for severely illness/accident-affected households becomes as good as zero. In contrast, for FIES, there is a tendency of a positive heterogenous treatment effect (p=.12). For FCS, the heterogenous treatment effect is as good as

[56] Information not displayed.
[57] The shock dummies are 1 if the respondents stated to have experienced a certain shock *"highly severe"* or *"alarmingly severe"* in the second period.
[58] Because of the double grouping, a PSM approach is not applied here.

zero. Also, theft, which is also a rather idiosyncratic shock compared to the others, indicates smaller tendencies of a .5-points-decrease in MDD (p=.16), while the other outcomes do not indicate clear differences. This leads to the conclusion that idiosyncratic shocks are likely to absorb the positive ATT on dietary diversity, measured by MDD.

Table 17: Heterogenous Treatment Effects of Adverse Shocks (DiDiD Estimators)

	Inverse FIES	FCS	MDD
Drought	-.108 (.775)	-3.230 (3.327)	.879 (.664)
Floods	-.071 (.348)	.006 (1.852)	-.158 (.301)
High agr. Costs	-.193 (.348)	.277 (2.092)	-.131 (.313)
High crops pests	.114 (.340)	3.318* (1.942)	.162 (.291)
Illness/accident	.627 (.399)	-.406 (2.457)	-1.116*** (.378)
Conflict	-.091 (.536)	1.016 (3.341)	-.390 (.450)
Theft	.364 (.476)	-2.388 (2.758)	-.489 (.345)

Notes: N=1,694, DiDiD estimators from Pooled OLS regressions for each shock separately, detailed regressions in Table A7 in the appendix. SE in parentheses and clustered on household level, * p < .1, ** p < .05, *** p < .01

There is a positive heterogenous treatment effect with *"high crop diseases"* in FCS by 3.3 points (p=0.088). The effect is comparatively large. Therefore, severely crop pest-affected beneficiary households perform significantly better in FCS than other treatment households. Including the control for support from other organizations during the intervention phase into this regression shows a large impact of other support with an increase in FCS by 2 points (p=.04)[59]. This indicates that potentially other organizations supported those who are highly affected by crop pests. However, Table A7 in the appendix highlights that being severely exposed to this crop diseases significantly decreases the control groups' FCS *("shock*wave")* (p<.1). In this work, his hypothesis will not be further investigated

Flood experience is of special interest in this work as the cash transfer was distributed to the severely flood-affected beneficiary households. No heterogenous treatment effect can be detected with the DiDiD, the estimators are as good as zero for all outcomes. This means that treatment households are similarly affected by the intervention, independent of their stated flood exposure. This supports the hypothesis that the cash transfer has offset the heterogenous treatment effect of the floods. However, it is crucial to consider that the subjective flood experience is not congruent to the assignment to cash distribution[60].

A possible explanation as to why there are no further heterogenous treatment effects is that covariate shocks affect a larger scale of people or even the whole sample. Therefore, there

[59] Information not displayed.
[60] The correlation within the TG between the two variables is -.08. Therefore, the assignment to the transfer is even more unlikely if the respondent stated a severe flood exposure.

might be less true variation in covariate shocks than it was stated in the survey. The overall negative trend in agricultural production in the control group hints to substantial challenges. Differences in ranking could be due to different perceptions. In this case, objective data would help to generate more conclusive results. The DiDiD analysis indicates that the interpretation of the ATT in section 4.2 are mainly stable. This supports the robustness of the findings[61].

4.4 Model Validity

4.4.1 Robustness Tests

There are two major threats to the internal validity of the conclusions. First, if the common trend assumption does not hold true, the counterfactual cannot be estimated reliably, so the ATT must be doubted. Second, if the SE are misreported, test statistics will be misleading, wherefore results might be interpreted wrong.

Table 18: Unrestricted Model including Attritted Households

	inverse FIES		FCS		MDD	
	R	UR	R	UR	R	UR
wave	.706***	.799***	-1.367	-1.471*	-.674***	-.582***
	(.129)	(.129)	(.714)	(.683)	(.114)	(.104)
treat	.384*	.425**	-1.382	-1.553*	-.974***	-.923***
	(.163)	(.148)	(.766)	(.694)	(.123)	(.108)
ATT	-.008	-.049	.918	1.089	1.135***	1.083***
	(.170)	(.170)	(.914)	(.883)	(.143)	(.135)
N	1694	1849	1694	1849	1694	1849

Notes: Pooled OLS with clustered SE on household level, R=Restricted Model, UR=Unrestricted model, both models without PSM extension, Standard errors in parentheses, * p < .1, ** p < .05, *** p < .01

A possible source of common trend disturbance is attrition. The attrition monitoring did not detect severe biases. To figure out if attrition might have disturbed the common trend, the unrestricted DiD model, including all available data from the baseline, can be compared to the regression results of restricted balanced panel[62]. Table 18 displays the results. First, the *"wave"* estimator indicates similar trends between the two models for all three indicators. Second, the differences between the control and treatment group (*"treat"*) emphasize that the group differences are roughly similar in both models for both outcomes.

[61] Table A7 in the appendix allows further interpretations of the relationship between shocks and the food security indicators, which are beyond the scope of this work.
[62] PSM is not applied here to detect the crude differences.

Third, the ATT is also not substantially different in the unrestricted models, so that both models lead to the same conclusions. All in all, the inclusion of all available baseline data does not change the interpretation. This supports my previous assumption of attrition being harmless. Even though, the unrestricted model is more efficient as it includes more data, it does not provide more conclusive indications.

Table 18 also allows to compare the restricted, unadjusted model with the PSM-adjusted model in the main analysis. The interpretations of the ATT on inverse FIES, FCS, and MDD remain the same. This stresses that the PSM approach did not add substantially more robustness to the model, either because the model was already very robust without the PSM approach or because the PSM could not improve the sample balance. In the main analysis model shows significant baseline differences after matching in inverse FIES and MDD. This suggests that the kernel-based PSM approach could not lead to perfect balance.

Another matching approach might be useful to sharpen the balance. Albeit, the survey participants live in the same region and very close to each other, a Moran's I test rejects the null hypothesis of spatial randomization for the outcome variables in the first wave ($p<0.01$ for all food security outcomes)[63] (Kondo, 2018). This stresses that accounting for these spatial differences might be a way to improve the comparability of the groups. A possible approach is to match the control households to the treatment households by PSM on the neighborhood level. Table 19 emphasizes that including neighborhoods into the PSM approach still leads to the same conclusions like in the main analysis for inverse FIES and MDD. For FCS, the positive tendency of an ATT turns into an insignificant ATT with a negative sign. Therefore, the interpretation of a positive tendency in ATT on FCS is not robust. The baseline differences indicate a slightly better balance in this model for FCS, as the difference is reduced to its half, and for inverse FIES, where the difference is now only significant at the 10% level instead of the 1% level. However, for the latter the difference magnitude remains the same. This model runs the risk of bad matches because the correlation between neighborhoods and food security outcomes is partially low (see Table A8 in the appendix[64]). Moreover, common support is doubted because as Figure A6 in the appendix indicates many observations that do not match, wherefore a substantial part of the information is lost[65]. Therefore, the significance reduction of the baseline difference in inverse FIES can be attributed to the lower sample size. The dropout leads to a total dropout, including attrition, of 30%. Especially in the beneficiary group, the increase in attrition

[63] The Moran's I test evaluates the degree of spatial correlation using GPS data (Kondo, 2018). The threshold distance for the spatial weights is set on 1 km. The distance within the intervention area is ~18.7 km from North to South and ~5.4 km from East to West.
[64] Critical neighborhoods are highlighted in red.
[65] Eight out of thirty-one neighborhoods are omitted from the estimation because these only contain observations from one group. After matching, 279 out of 384 control observations and 404 of the 463 treatment observations are included.

from 7% to 13% disturbs the randomization of the treatment group sampling, whereas the full average treatment effect of the treated is assessed worse than in the main model. The weighting might cause a more harmful bias than it corrects, whereas the model without spatial PSM matching is preferred[66].

Table 19: Treatment and Food Security with Spatial PSM

	Baseline			Endline			ATT
	CG	TG	Diff	CG	TG	Diff	
Inverse FIES	2.264	2.663	.399*	2.741	3.317	.576**	.177
			(.221)			(.236)	(.224)
FCS	28.641	28.046	-.595	28.076	27.052	-1.024	-.429
			(1.099)			(1.216)	(1.473)
MDD	3.709	3.007	-.701***	3.130	3.406	-.276*	.977***
			(.160)			(.158)	(.210)

Notes: N=1,366, CG=Control Group, TG=Treatment Group, Diff=Group Difference, Pooled OLS regressions, Kernel-based PSM, SE in parentheses and clustered on household level, * p < .1, ** p < .05, *** p < .01, outcome variables in first column

Next, I focus on the goodness of the standard errors. In this analysis, a pooled OLS model was used, even though a fixed effects model is also applicable. The models differ in SE. To control if the test statistics change substantially with another model approach, Table 20 displays the ATTs of the applied model and of a FE model (columns (1) and (2))[67]. The SE in parenthesis emphasize that the SE are nearly equal in both models, the interpretation of the test statistics does not differ. Therefore, FE does not provide any advantage compared to the applied model.

Table 20: Standard Error Evaluation of the ATT

	(1)	(2)	(3)
	Pooled OLS	FE	Two-way clustered SE
Inverse FIES	-.008 (.17045)	-.008 (.17040)	-.008 (.16543)
FCS	.918 (.91387)	.918 (91360)	.918 (1.21617)
MDD	1.135 (.14310)	1.135 (.14306)	1.135 (.18746)

Notes: SE in parenthesis and clustered on household level. (1): Pooled OLS model from main analysis without kernel-based PSM approach; (2): Fixed Effects model; (3) Pooled OLS model with SE additionally clustered on neighborhood level

[66] Furthermore, GPS locations are assessed from the interview place and not necessarily from the residency location.
[67] The applied model is displayed without the kernel-based PSM approach so that the coefficients are equal to the model alternatives. Like this, the SE can be compared more precisely.

Ignoring group correlations violates the OLS assumption of zero conditional mean and leads to an overestimation of test statistics. SE in the main analysis model are not clustered locally. This can result in mistaken coefficient interpretation. Column (3) in Table 20 shows a two-way clustered model that allows within-neighborhood correlation combined to the clustering on household level. The SE increase slightly for FCS and MDD. Nevertheless, the SE do not change enough to alter the interpretation of the test statistics. Additional neighborhood clustering leads to the same conclusions like in 4.1, wherefore, this extension is redundant in this model[68].

In conclusion, these robustness checks imply a high precision of the model and highlight that a range of model adjustments imply the same conclusions. Therefore, the internal validity is robust, so that the results can be attributed causally to the program. The standard errors are very similar under different approaches. This emphasizes the adequacy of the interpretation. An additional adjustment for heteroskedasticity is not necessary because the clustered standard error already apply heteroskedasticity robust SE. Likewise, the common trend assumption appears very stable. For a stronger validity of the common trend, multiple data waves before treatment would help to evaluate the group development. Also, objective shock data could rule out heterogenous treatment effects through adverse shocks more reliably. A stronger internal validity could be reached with a RCT.

4.4.2 External Validity

Fiorella et al. (2016) stressed that conclusions on agricultural interventions are difficult to generalize. This study is a case study. Therefore, to apply the conclusions from this analysis to other settings, one must consider that the present data are only representative for a certain group of people, which are mainly illiterate and poor farmers. Also, the assessment only covers short-term impacts. The intervention design includes a range of intervention tools, wherefore the impact cannot be differentiated for the single components, other than the cash transfer. Therefore, for a valid adaption of conclusions, the programs should be similar. Furthermore, the country context is very specific, as South Sudan is among the least developed countries and the country is highly affected by conflict, with high prevalence of adverse shocks. Particularly in conflict-exposure, this study contrasts from the previous literature. If researchers ensure the comparability in intervention design and target groups, these conclusions are transferable to other regions in South Sudan as well as to more countries with a similar context in Sub-Saharan Africa, e.g., Burkina Faso, Central

[68] It was already argued in 3.2.3 that misspecified clustering can be harmful. Other clustering approaches are not feasible in this setting due to data limitations.

African Republic, Chad, Democratic Republic of the Congo, Ethiopia, Mali, Mozambique, Niger, or Somalia[69].

5 Discussion

This analysis showed that in the short run, the program development goal of an enhanced production could be reached. That the production of the treatment households remained rather stable compared to the control households in a very challenging production year, indicates that beneficiary households became more resilient towards adverse shocks and other challenges. However, a robust resilience improvement remains to be investigated in future research. This is in line with other research from peaceful, low-developed settings, which frequently assess increases of agricultural outputs. Therefore, this thesis emphasizes that small-scale agricultural interventions have the potential to empower agricultural production in a conflict-affected least developed setting with a high adverse shock exposure.

In contrast to most of the evidence from peaceful settings, the food security indicators did not improve substantially apart from MDD. However, the lack of assessed impacts can have several reasons unconnected to conflict. First, the evaluation was too early to assess clear changes because the program impacts in the agricultural context are assumed to be lagged due to learning effects and plant growth periods. Therefore, the program impact is assumed to be stronger at a later stage.

Second, the year 2020 was particularly shaded by crises additional to the high degree of violence. On the one hand side, the Covid-19 pandemic challenged markets and disturbed the normal processes, on the other hand side, the severe flood on the study-site restricted the agricultural potential and therefore, probably attenuated the real program impact. As the program runs for further two years after the assessment of the second wave, future evaluations will be informative about the stability and the sustainability of the assessed impacts. However, also the first half of 2021 is permeated by crisis. The markets are still challenged by the Covid-19 pandemic and violence indicates an increasing trend; nevertheless, other severe adverse shocks comparable to the flood in 2020 are not observed (World Bank, 2021c). Therefore, the high exposure to adverse shocks has to be seen a permanent condition. Comparing the results to assessments in less fragile settings, the analysis implicates that the complex context might attenuate the program impact on food security.

Third, the indicators might not assess the full impact on food security. Contemporary science does not provide a feasible measure that covers all dimensions of food insecurity;

[69] These countries show the following similarities to South Sudan: location in Sub-Saharan Africa, high or medium-intense conflict (World Bank, 2021b), and classified as least developed countries (UNCTAD, n.d.). Ethiopia is added to the list because of the recent escalation in the Tigray region.

therefore, the applied indicators are limited to assess the full extent of the households' food security (Brück et al., 2018). Further, the measures are not sufficiently sensible towards small improvements. This especially applies for FIES, which covers a whole year and therefore also includes the time, when the program was just implemented and could not influence food security yet. A narrower indicator like the HFIAS would be more informative. Alternatively, FIES could be assessed more frequently[70].

Fourth, the intervention impact on production might have been too small to influence FCS and inverse FIES substantially. The intervention had a clear positive impact on production, however, the analysis also revealed that increases in the food security indicators are associated with substantially larger production increases than the assessed increases.

In the end, all indicators show some improvements in the breakdown analysis. While in FIES, the likelihood for respondents to experience the most critical situation of not eating anything for a whole day decreased significantly, FCS reveals adjustments in household food choices through the treatment. Furthermore, negative developments in FCS, MDD, and production are assessed for the control group, while the treatment group rather remained stable or improved its outcomes. Therefore, a positive impact can be attributed to the intervention. The positive and significant treatment effect on MDD implies that the intervention could contribute to an improved dietary diversity, which, in turn, might lower the likelihood of malnutrition. Especially for children, the prevention of malnutrition has long-run impacts for their future that are not assessable in this study yet.

The three food security indicators differ in interpretations. This can have several explanations. First, the indicators cover different recall times. Second, the scores assess different dimensions, while FIES assesses a broader insight covering constraints to food access but also psychological spheres such as anxiety, FCS and MDD cover nutritional diversity with different groupings. Still, that MDD and FCS indicate different interpretations is noteworthy. The two indicators differ because MDD assesses the respondents' food security while FCS focuses on the whole household. This implies intra-household differences in dietary diversity. Apparently, the survey respondents benefit more from the program than the whole household. Furthermore, the grouping of the food types differs substantially.

The analysis indicates that the positive impact on MDD is rather attributed to the combined treatment, consisting of the cash transfer and SAI, even though the cash transfer did not induce increases in food spending. This leads to the assumption that the transfer was probably partially spent on agricultural investments like on fertilizers. The treatment group's agricultural output increased substantially through treatment. Therefore, the intervention impact can be attributed to the combined intervention; however, to explain the full impact

[70] This would increase the assessment costs.

on MDD, the increase in agricultural productivity is too small. For inverse FIES and FCS, the analysis did not reveal any striking differences in outcomes between the treatment households that received the cash transfer and those that did not. The people who received this transfer were affected by the flooding and are not substantially worse off compared to the other treatment households. This implies that the transfer could compensate the negative flood impact. Therefore, the transfer might have an impact on food security that is not assessable in this evaluation. The transfer was distributed three-four months before the assessment of the second wave, so that the money was possibly already spent on food, which would affect FIES, but was possibly not regarded in the expenditure assessment. Additionally, the sample size of the treatment group that did not receive the cash transfer is not high enough for powerful conclusions. Therefore, for a robust assessment, more observations from treatment households that did not receive the transfer are needed. This transfer has to be treated rather as emergency aid, so that these implications cannot be compared to regular transfers, subsidies or vouchers. Regular transfers are more likely to be spent directly on food compared to one time payments that are rather invested (Haushofer and Shapiro, 2017, 2018). In conclusion, the cash transfer combined to the SAI appears to be a valuable approach to mitigate shock impacts.

The exposure to adverse shocks did mainly not induce any heterogenous treatment effects. Especially, covariate shocks do not indicate differences in impact size[71]. This does not mean that these shocks do not attenuate the treatment effect. The households are permanently exposed to adverse shocks. That there are no assessable differences between treatment households that experienced covariate adverse shocks severely and treatment households that did not, may have two major reasons. First, it is plausible that all households are affected by a comparable extent by the same shocks in the same year and the differences in assessments emerge from different perceptions. Therefore, the shock impact is already assessed in the trend development. Second, some adverse shocks induce each other, such as severe catastrophes that induce high losses like a flood could increase the level of conflict or induce thefts. Many households experienced a sum of adverse shocks, wherefore the impact of one single shock cannot be differentiated clearly from the others. For a more robust assessment of the covariate adverse shock impacts, more data from other regions with a different shock exposure would be valuable. The most informative data source would be objective shock data, which could be captured by satellite data that tracks the extent of a flood, for example.

For the idiosyncratic shock of an illness or an accident of the household head, the implications are clearer. The occurrence of such a challenge absorbs the whole positive impact on

[71] Excluding the positive heterogenous treatment effect of high crop pests, which was discussed in detailed in section 4.3.

MDD. This is reasonable because such events frequently imply a high financial burden, the absence of one labor force, and a psychological burden, which are not reflected in the whole population average. Furthermore, emergency support is rather provided to victims of covariate events. This emphasizes that large scale aid is limited to take individual impairing events into account.

By its end, this work contributes to further research opportunities. First, a cost-benefit analysis quantifying the benefits monetarily, for example, by disability-adjusted life years, would be informative about the program's efficiency. Second, long-term implications of the program could be assessed after the finalization of the program. This is crucial knowledge for an adequate evaluation of the program's sustainability. Third, it is also of special interest to assess the regional spillover effects as the program affects regional economic development. A possible approach could be to sample a control group living in the same neighborhood as the treatment group and additionally sample a control group that lives further away from the beneficiaries to disentangle the spillovers. In this manner, the full program impact may be assessed ever more precisely. Fourth, these data also allow to differentiate between groups and assess heterogenous treatment effects for subgroups. Possible subgroups might contain female household heads, widowed household heads, IDP households, or farming households with or without land access to water. Fifth, the inclusion of other outcomes that assess further dimensions of food security could be informative about the full program impact. For example, the measurement of hemoglobin allows for the detection of whether or not the agricultural intervention could increase the iron intake and, therefore, prevent health issues such as anemia. Other examples are the measurement of the weight and height of children to assess the treatment effect on anthropometric outcomes, the measurement of the caloric food intake or the assessment of HFIAS. Lastly, the program is also expected to have impacts on resilience towards adverse shocks, health, peace, women empowerment, economic development, and wellbeing, which would also be interesting to be investigated.

6 Conclusion

In conclusion, this study emphasizes that the agricultural intervention, consisting of capacity building, in-kind provisions, and cash transfers to flood-affected households increased the dietary composition of the participants by one additional component per day one year after the program implementation. MDD and FCS indicate changes in food choices away from staples towards more pulses, nuts, fruits, certain types of vegetables, meat, and eggs. Still, the average treatment household remains below the threshold of micronutrient adequacy. The analysis stresses that the consumption of green, leafy vegetables decreased, and the intake of sugar increased through the intervention. This is a rather unintended outcome. On the Food Insecurity Experience Scale, the intervention impact is as good as zero. However, the likelihood to experience a whole day without eating decreased through treatment. On the Food Consumption Score, the impact is small and insignificant, even though clear compositional changes in consumption are detected. Robustness tests indicate that the impact on FCS is questionable.

Comparing the treatment components, the analysis indicates that the composite treatment has a no substantial impact on food security in absence of cash transfers and high flood-exposure in the treatment group like the main analysis for FCS and inverse FIES. The cash transfer did not lead to increased spending on food, wherefore it is assumed that this emergency aid was rather used to compensate the flood damage. However, the positive program impact on MDD is rather attributed to the combined intervention. Therefore, cash distributions successfully complement SAI and compensate emergency losses, whereas it is reasonable to extent in-kind assistance with direct transfers in this setting. In line with most of the previous research, the results show that the agricultural support has a significant positive impact on the harvest of different crops. These are, in turn, associated with better food security outcomes. An increased harvest shows a mostly positive and significant association with food security measures. However, to increase the outcomes substantially, large increases in harvest are needed.

The analysis implies that a perceived severe adverse shock experience mainly did not have a heterogenous treatment effect on food security. A severe impairment of the household head absorbed the most part of the positive treatment effect in MDD, while a high exposure to crop pests is associated with an increase of FCS. This is likely to be explained by other organizations focusing on this threat.

All in all, the impact evaluation suggests that small-scale agricultural interventions can improve dietary diversity in highly conflict- and shock-affected least developed settings in the short run, however, the impact magnitude indicates that the program can only be understood as a part of the solution rather than as the solution itself. The treatment impacts will potentially be more pronounced at the end of the intervention in three years. A more striking impact is expected particularly if the ongoing conflicts, the Covid-19 pandemic, and the flood consequences attenuate in the following years.

References

Abadie, A., Athey, S., Imbens, G., Wooldridge, J., 2017. When Should You Adjust Standard Errors for Clustering? (No. w24003). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w24003

Abadie, A., Imbens, G.W., 2006. Large Sample Properties of Matching Estimators for Average Treatment Effects. Econometrica 74, 235–267. https://doi.org/10.1111/j.1468-0262.2006.00655.x

Abraham, M., Pingali, P., 2020. Transforming Smallholder Agriculture to Achieve the SDGs, in: Gomez y Paloma, S., Riesgo, L., Louhichi, K. (Eds.), The Role of Smallholder Farms in Food and Nutrition Security. Springer International Publishing, Cham, pp. 173–209. https://doi.org/10.1007/978-3-030-42148-9_9

ACLED, n.d. Dashboard. URL: https://acleddata.com/dashboard/#/dashboard. (Accessed 16.07.21).

Ajak, B.J., 2018. Livelihood resilience and adaptation to climate variability in post-conflict South Sudan-A review. AG 39. https://doi.org/10.18805/ag.R-1788

Aker, J.C., 2017. Comparing Cash and Voucher Transfers in a Humanitarian Context: Evidence from the Democratic Republic of Congo. The World Bank Economic Review 31, 44–70. https://doi.org/doi: 10.1093/wber/lhv055

Aker, J.C., 2013. Cash or Coupons? Testing the Impacts of Cash versus Vouchers in the Democratic Republic of Congo. SSRN Journal. https://doi.org/10.2139/ssrn.2241237

Aker, J.C., Boumnijel, R., McClelland, A., Tierney, N., 2011. Zap It to Me: The Short-Term Impacts of a Mobile Cash Transfer Program. SSRN Journal. https://doi.org/10.2139/ssrn.1931641

Amnesty International, 2020. South Sudan: United Nations arms embargo must remain in place after surge in violence against civilians in 2020, Advocacy Brief.

Andersen, L., Cardona, M., 2013. Building Resilience against Adverse Shocks: What are the determinants of vulnerability and resilience? Institute for Advanced Development Studies.

Angrist, J., Pischke, J.-S., 2015., Chapter 5: Differences-in-Differences. in: Mastering 'metrics: The path from cause to effect. Princeton University Press, pp. 178–208.

Angrist, J.D., Pischke, J.-S., 2008. Chapter 8: Nonstandard Standard Error Issues, in: Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton, pp. 221–244.

Arias, M.A., Ibáñez, A.M., Zambrano, A., 2019. Agricultural production amid conflict: Separating the effects of conflict into shocks and uncertainty. World Development 119, 165–184. https://doi.org/10.1016/j.worlddev.2017.11.011

Aurino, E., Giunti, S., 2021. Social Protection for Child Development in Crisis: A Review of Evidence and Knowledge Gaps. The World Bank Research Observer lkab007. https://doi.org/10.1093/wbro/lkab007

Bageant, E., Liu, Y., Diao, X., 2016. Agriculture-Nutrition Linkages and Child Health in the Presence of Conflict in Nepal. Markets, Trade and Institutions Division - Development Strategy and Governance Division IFPRI Discussion Paper.

Baliki, G., Brück, T., Schreinemachers, P., Uddin, Md.N., 2019. Long-term behavioural impact of an integrated home garden intervention: evidence from Bangladesh. Food Sec. 11, 1217–1230. https://doi.org/10.1007/s12571-019-00969-0

Ballard, T., Coates, J., Swindale, A., Deitchler, M., 2011. Household Hunger Scale: Indicator Definition and Measurement Guide. USAID.

Banerjee, A., Duflo, E., 2011. Chapter 6: Barefoot Hedge-Fund Managers, in: Poor Economics - A Radical Rethinking of the Way to Fight Global Poverty. Public Affairs, New York, pp. 247–288.

Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Pariente, W., Shapiro, J., Thuysbaert, B., Udry, C., 2015. A multifaceted program causes lasting progress for the very poor: Evidence from six countries. Science 348, 1260799–1260799. https://doi.org/10.1126/science.1260799

Berti, P.R., Krasevec, J., FitzGerald, S., 2004. A review of the effectiveness of agriculture interventions in improving nutrition outcomes. Public Health Nutr. 7, 599–609. https://doi.org/10.1079/PHN2003595

Bertrand, M., Duflo, E., Mullainathan, S., 2004. How Much Should We Trust Differences-in-Differences Estimates? The Quarterly Journal of Economics 119, 249–275.

Bizikova, L., Jungcurt, S., McDougal, K., Tyler, S., 2020. How can agricultural interventions enhance contribution to food security and SDG 2.1? Global Food Security 26, 100450. https://doi.org/10.1016/j.gfs.2020.100450

Blackorby, C., Donaldson, D., 1988. Cash versus Kind, Self-Selection, and Efficient Transfers. The American Economic Review 78, 691–700.

Bozzoli, C., Brück, T., 2009. Agriculture, Poverty, and Postwar Reconstruction: Micro-Level Evidence from Northern Mozambique. Journal of Peace Research 46, 377–397. https://doi.org/10.1177/0022343309102658

Brück, T., d'Errico, M., 2019. Food security and violent conflict: Introduction to the special issue. World Development 117, 167–171. 10.1016/j.worlddev.2019.01.007

Brück, T., Díaz Botía, O.M., Ferguson, N.T.N., Ouédraogo, J., Ziegelhöfer, Z., 2019. Assets for Alimentation? The Nutritional Impact of Assets-based Programming in Niger. The Journal of Development Studies 55, 55–74. https://doi.org/10.1080/00220388.2019.1687876

Brück, T., Justino, P., Verwimp, P., Avdeenko, A., Tedesco, A., 2015. Measuring Violent Conflict in Micro-level Surveys: Current Practices and Methodological Challenges. World Bank Res Obs lkv011. https://doi.org/10.1093/wbro/lkv011

Brüderl, J., 2010. Kausalanalyse mit Paneldaten, in: Wolf, C., Best, H. (Eds.), Handbuch der sozialwissenschaftlichen Datenanalyse. VS Verlag für Sozialwissenschaften, Wiesbaden, pp. 963–994. https://doi.org/10.1007/978-3-531-92038-2_36

Cafiero, C., Viviani, S., Nord, M., 2018. Food security measurement in a global context: The food insecurity experience scale. Measurement 116, 146–152. https://doi.org/10.1016/j.measurement.2017.10.065

Cameron, A.C., Miller, D.L., 2015. A Practitioner's Guide to Cluster- Robust Inference. The Journal of Human Resources 50, 317–372.

Cameron, A.C., Trivedi, P.K., 2005a. Chapter 3: Microeconomic Data Structures, in: Microeconometrics: Methods and Applications. Cambridge University Press, Cambridge; New York, pp. 39–64.

Cameron, A.C., Trivedi, P.K., 2005b. Chapter 25: Treatment Evaluation, in: Microeconometrics: Methods and Applications. Cambridge University Press, Cambridge; New York, pp. 860–898.

Cameron, A.C., Trivedi, P.K., 2005c. Chapter 22: Linear Panel Models: Entensions, in: Microeconometrics: Methods and Applications. Cambridge University Press, Cambridge; New York, pp. 743–778.

Cameron, A.C., Trivedi, P.K., 2005d. Chapter 21: Linear Panel Models: Basics, in: Microeconometrics: Methods and Applications. Cambridge University Press, Cambridge; New York, pp. 697–742.

Carletto, G., Ruel, M., Winters, P., Zezza, A., 2015. Farm-Level Pathways to Improved Nutritional Status: Introduction to the Special Issue. The Journal of Development Studies 51, 945–957. https://doi.org/10.1080/00220388.2015.1018908

Center for Military History and Social Sciences of the Bundeswehr, 2018. Wegweiser der Geschichte - Sudan und Südsudan. Schöningh.

Cesarini, D., Lindqvist, E., Notowidigdo, M.J., Östling, R., 2017. The Effect of Wealth on Individual and Household Labor Supply: Evidence from Swedish Lotteries. American Economic Review 107, 3917–3946. https://doi.org/10.1257/aer.20151589

CLiMIS South Sudan, 2021. Minimum Expenditure Basket. South Sudan National Bureau of Statistics and Ministry of Agriculture and Food Security.

CLiMIS South Sudan, n.d.a. Cereal Data Trend Analysis. South Sudan National Bureau of Statistics and Ministry of Agriculture and Food Security.

CLiMIS South Sudan, n.d.b. Population Summaries: South Sudan Technical Working Group. URL https://climis-southsudan.org/ipc/pop_summaries_county

Coates, J., Swindale, A., Bilinsky, P., 2007. Household Food Insecurity Access Scale (HFIAS) for Measurement of Household Food Access: Indicator Guide (v. 3). Food and Nutrition Technical Assistance Project, Academy for Educational Development, Washington, DC.

Columb, M.O., Stevens, A., 2008. Power analysis and sample size calculations. Current Anaesthesia & Critical Care 19, 12–14. https://doi.org/10.1016/j.cacc.2007.03.011

Cunha, J.M., 2014. Testing Paternalism: Cash versus In-Kind Transfers. American Economic Journal: Applied Economics 6, 195–230. https://doi.org/10.1257/app.6.2.195

Cuong, N.V., 2012. Selection of Control Variables in Propensity Score Matching: Evidence from a Simulation Study. MPRA Paper.

CurrencyRate, 2021. 1 EUR Euro to SSP South Sudanese Pound. URL https://eur.currencyrate.today/ssp

d'Agostino, G., Pieroni, L., Scarlata, M., 2013. Social Protection and Food Security in Sub-Saharan Africa: An Evaluation of Cash Transfer Programmes. Munich Personal RePEc Archive 49536.

d'Errico, M., Ngesa, O., Pietrelli, R., 2021. Assistance in chronic conflict areas: evidence from South Sudan. Journal of Development Effectiveness 13, 145–165. https://doi.org/10.1080/19439342.2021.1924835

Dalberg, 2012. Catalysing Smallholder Agricultural Finance, Dalberg Global Development Advisors.

Deaton, A., 1991. Saving and Liquidity Constraints. Econometrica 59, 1221. https://doi.org/10.2307/2938366

Dercon, S., 2004. Risk, Insurance, and Poverty: A Review, in: Dercon, S. (Ed.), Insurance Against Poverty. Oxford University Press, pp. 9–37. https://doi.org/10.1093/0199276838.003.0002

Doss, C., Oduro, A.D., Deere, C.D., Swaminathan, H., Baah-Boateng, W., J. Y., S., 2018. Assets and shocks: a gendered analysis of Ecuador, Ghana and Karnataka, India. Canadian Journal of Development Studies / Revue canadienne d'études du développement 39, 1–18. https://doi.org/10.1080/02255189.2017.1316244

Duflo, E., Glennerster, R., Kremer, M., 2007. Chapter 61 Using Randomization in Development Economics Research: A Toolkit, in: Handbook of Development Economics. Elsevier, pp. 3895–3962. https://doi.org/10.1016/S1573-4471(07)04061-2

Duncan, B., Joey, G., Isabel, L., David, M., D, H., Derek, 2020. Strengthening smallholder agriculture is essential to defend food and nutrition security and rural livelihoods in Myanmar against the COVID-19 threat: Elements for a proactive response. Intl Food Policy Res Inst.

Ellis, F., Maliro, D., 2013. Fertiliser Subsidies and Social Cash Transfers as Complementary or Competing Instruments for Reducing Vulnerability to Hunger: The Case of Malawi. Development Policy Review 31, 575–596. https://doi.org/10.1111/dpr.12026

Esarey, J., Menger, A., 2019. Practical and Effective Approaches to Dealing With Clustered Data. PSRM 7, 541–559. https://doi.org/10.1017/psrm.2017.42

FAO, 2019. South Sudan - Resilience Strategy 2019-2020. Food and Agriculture Organization of the United Nations, Rome.

FAO, 2017. Nutrition-sensitive agriculture and food systems in practice - Options for intervention. Food and Agriculture Organization of the United Nations, Rome.

FAO, 1996. The Rome Declaration on World Food Security, The World Food Summit. Food and Agriculture Organization of the United Nations, Rome.

FAO, n.d. Analyse the Data | Voices of the Hungry | Food and Agriculture Organization of the United Nations [WWW Document]. Food and Agriculture Organization of the United Nations. URL http://www.fao.org/in-action/voices-of-the-hungry/analyse-data/en/ (accessed 2.6.21).

FAO, FHI, 2016. Minimum Dietary Diversity for Women- A Guide to Measurement. Rome.

FAO, IFAD, UNICEF, WFP, WHO, 2020. The State of Food Security and Nutrition in the World 2020. Transforming food systems for affordable healthy diets. Food and Agriculture Organization of the United Nations, Rome. https://doi.org/10.4060/ca9692en

Fiorella, K.J., Chen, R.L., Milner, E.M., Fernald, L.C.H., 2016. Agricultural interventions for improved nutrition: A review of livelihood and environmental dimensions. Global Food Security 8, 39–47. https://doi.org/10.1016/j.gfs.2016.03.003

Fredriksson, A., Oliveira, G.M. de, 2019. Impact evaluation using Difference-in-Differences. RAUSP 54, 519–532. https://doi.org/10.1108/RAUSP-05-2019-0112

FSIN, 2021. Global Report on Food Crisis - Joint Analysis for better Decisions. Global Network against Food Crisis.

FSIN, 2020. Global Report on Food Crises - Joint Analysis for better Decisions. Global Network against Food Crisis.

Galhena, D.H., Freed, R., Maredia, K.M., 2013. Home gardens: a promising approach to enhance household food security and wellbeing. Agriculture & Food Security 2, 8.

Gao, J., Bradford, M., 2018. Weather shocks, coping strategies, and consumption dynamics in rural Ethiopia. World Development 101, 268–283.

Gitter, S.R., Manley, J., Bernstein, J., Winters, P., 2017. Do Agricultural Support and Cash Transfer Programmes Improve Nutritional Status? SSRN Journal. https://doi.org/10.2139/ssrn.3307396

Günther, I., Harttgen, K., 2009. Estimating Households Vulnerability to Idiosyncratic and Covariate Shocks: A Novel Method Applied in Madagascar. World Development 37, 1222–1234. https://doi.org/10.1016/j.worlddev.2008.11.006

Haushofer, J., Shapiro, J., 2018. The Long-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya.

Haushofer, J., Shapiro, J., 2017. The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya. The Quarterly Journal of Economics 132, 2057–2060. https://doi.org/10.1093/qje/qjx039

HDX, 2021. Prices of food commodities [WWW Document]. Humanitarian Data Exchange. URL https://data.humdata.org/hxlpreview/show;url=https:%2F%2Fdata.humdata.org%2Fdataset%2F4f48c45b-1a20-4478-b62d-bf637b8d4666%2Fresource%2F457a4550-ba5c-4858-84ca-5ce8d52bae5c%2Fdownload%2Fwfp_food_median_prices_south-sudan.csv;embeddedSource=World%2520Food (accessed 18.5.21).

Heinrich, C., Maffioli, A., Vázquez, G., 2010. A Primer for Applying Propensity-Score Matching- Impact-Evaluation Guidelines. Inter-American Development Bank.

Heltberg, R., Lund, N., 2009. Shocks, Coping, and Outcomes for Pakistan's Poor: Health Risks Predominate. The Journal of Development Studies 45, 889–910. https://doi.org/10.1080/00220380902802214

Hidrobo, M., Hoddinott, J., Peterman, A., Margolies, A., Moreira, V., 2014. Cash, food, or vouchers? Evidence from a randomized experiment in northern Ecuador. Journal of Development Economics 107, 144–156. https://doi.org/10.1016/j.jdeveco.2013.11.009

Hsiao, C., 2007. Panel data analysis—advantages and challenges. TEST 16, 1–22. https://doi.org/10.1007/s11749-007-0046-x

IFAD, UNEP, 2013. Smallholders, food security, and the environment. International Fund for Agricultural Development and United Nations Environment Programme.

Imai, K., Kim, I.S., 2020. On the Use of Two-Way Fixed Effects Regression Models for Causal Inference with Panel Data. Polit. Anal. 1–11. https://doi.org/10.1017/pan.2020.33

IMF, 2021. Südsudan: Bruttoinlandsprodukt (BIP) pro Kopf in jeweiligen Preisen Schätzungen und Prognosen von 2011 bis 2026 (in US-Dollar) [Graph]. Statista. URL https://de.statista.com/statistik/daten/studie/387076/umfrage/bruttoinlandsprodukt-bip-pro-kopf-im-suedsudan/ (accessed 17.5.21).

IPC Global Partners, 2020. South Sudan: Acute Food Insecurity and Acute Malnutrition Situation for October 2020 – July 2021. Integrated Food Security Phase Classification.

IPC Global Partners, 2019. Technical Manual Version 3.0 - Evidence and Standards for Better Food Security and Nutrition Decisions. Food and Agriculture Organization of the United Nations, Rome.

IPC Global Partners, n.d. South Sudan: Consolidated Findings from the IPC Technical Working Group and External Reviews. Food and Agriculture Organization of the United Nations.

Jayne, T.S., Mather, D., Mghenyi, E., 2010. Principal Challenges Confronting Smallholder Agriculture in Sub-Saharan Africa. World Development 38, 1384–1398. https://doi.org/10.1016/j.worlddev.2010.06.002

John Hopkins University, 2021. South Sudan - COVID-19 Overview - Johns Hopkins [WWW Document]. Johns Hopkins Coronavirus Resource Center. URL https://coronavirus.jhu.edu/region/south-sudan (accessed 18.6.21).

Jok, J.M., 2021. Lessons in Failure: Peacebuilding in Sudan/South Sudan, in: McNamee, T., Muyangwa, M. (Eds.), The State of Peacebuilding in Africa. Springer International Publishing, Cham, pp. 363–377. https://doi.org/10.1007/978-3-030-46636-7_20

Keeley, B., 2012. What is aid?, in: From Aid to Development, OECD Insights. OECD, pp. 46–65. https://doi.org/10.1787/9789264123571-4-en

Kondo, K., 2018. Testing for global spatial autocorrelation in Stata. Research Institute of Economy, Trade, and Industry.

Larsen, A.F., Lilleør, H.B., 2014. Beyond the Field: The Impact of Farmer Field Schools on Food Security and Poverty Alleviation. World Development 64, 843–859. https://doi.org/10.1016/j.worlddev.2014.07.003

Lechner, M., Rodriguez-Planas, N., Fernández Kranz, D., 2015. Difference-in-Difference Estimation by FE and OLS when there is Panel Non-Response. IZA Discussion Paper.

Levin, C.E., Long, J., Simler, K.R., Johnson-Welch, C., 2003. Cultivating nutrition: a survey of viewpoints on integrating agriculture and nutrition. IFPRI 154.

Lowder, S.K., Skoet, J., Raney, T., 2016. The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. World Development 87, 16–29. https://doi.org/10.1016/j.worlddev.2015.10.041

MacKinnon, J.G., Webb, M.D., 2017. Pitfalls when Estimating Treatment Effects Using Clustered Data. Queen's Economics Department Working Paper 1387.

Macrotrends LLC, 2020. Crude Oil Prices - 70 Year Historical Chart. URL https://www.macrotrends.net/1369/crude-oil-price-history-chart (accessed 12.9.20).

Margolies, A., 2019. Seeding change? Improving and understanding household food security in a nutrition-sensitve agriculture program in Malawi: A mixed-methods study. Johns Hopkins University 181.

Marquis, G.S., Colecraft, E.K., Kanlisi, R., Aidam, B.A., Atuobi-Yeboah, A., Pinto, C., Aryeetey, R., 2018. An agriculture-nutrition intervention improved children's diet and growth in a randomized trial in Ghana. Matern Child Nutr 14, e12677. https://doi.org/10.1111/mcn.12677

Martin-Shields, C.P., Stojetz, W., 2019. Food security and conflict: Empirical challenges and future opportunities for research and policy making on food security and conflict. World Development 119, 150–164.

Mary, S., Shaw, K., Colen, L., Gomez y Paloma, S., 2020. Does agricultural aid reduce child stunting? World Development 130, 104951. https://doi.org/10.1016/j.worlddev.2020.104951

Mataf, 2021. History of daily rates SSP /EUR since Sunday, 15 November 2015. URL https://www.mataf.net/en/currency/converter-SSP-EUR/data (accessed 30.4.21).

Maxwell, D., Vaitla, B., Coates, J., 2014. How do indicators of household food insecurity measure up? An empirical comparison from Ethiopia. Food Policy 47, 107–116. https://doi.org/10.1016/j.foodpol.2014.04.003

Maxwell, S., Smith, M., 1992. Household Food Security: A Conceptual Review - Part 1 Household Food Security: A Conceptual Review, 1–72.

Meyer, B.D., 1995. Natural and Quasi-Experiments in Economics. Journal of Business & Economic Statistics, Symposium on Program and Policy Evaluation. 13, 151–161.

Moyo, J.M., Bah, E., Verdìer-Chouchane, A., 2015. Chapter 2.1: Transforming Africa's agriculture to improve competitiveness, in: The Africa Competitiveness Report. pp. 37–52.

Nsabuwera, V., Hedt-Gauthier, B., Khogali, M., Edginton, M., Hinderaker, S.G., Nisingizwe, M.P., Tihabyona, J. de D., Sikubwabo, B., Sembagare, S., Habinshuti, A., Drobac, P., 2016. Making progress towards food security: evidence from an intervention in three rural districts of Rwanda. Public Health Nutr. 19, 1296–1304. https://doi.org/10.1017/S1368980015002207

Nyikahadzoi, K., Siziba, S., Mango, N., Mapfumo, P., Adekunhle, A., Fatunbi, O., 2012. Creating food self reliance among the smallholder farmers of eastern Zimbabwe: exploring the role of integrated agricultural research for development. Food Sec. 4, 647–656. https://doi.org/10.1007/s12571-012-0218-8

OCHA, 2021. South Sudan: Flooding Situation Report - Inter-Cluster Coordination Group. UN Office for the Coordination of Humanitarian Affairs.

OCHA, 2020. South Sudan - Flooding Snapshot. Office for the Coordination of Humanitarian Affairs of the United Nations.

Olney, D.K., Pedehombga, A., Ruel, M.T., Dillon, A., 2015. A 2-year integrated agriculture and nutrition and health behavior change communication program targeted to women in Burkina Faso reduces anemia, wasting, and diarrhea in children 3–12.9 months of age at baseline: a cluster-randomized controlled trial. The Journal of nutrition 145, 1317–1324.

Pan, Y., Smith, S.C., Sulaiman, M., 2018. Agricultural Extension and Technology Adoption for Food Security: Evidence from Uganda. American Journal of Agricultural Economics 100, 1012–1031. https://doi.org/10.1093/ajae/aay012

Puri, J., Aladysheva, A., Iversen, V., Ghorpade, Y., Brück, T., 2017. Can rigorous impact evaluations improve humanitarian assistance? Journal of Development Effectiveness 9, 519–542.

Quinn, C., Fox, A., Baroang, K., Evans, D., Gomes, M., Habib, J., 2019. South Sudan Climate Vulnerability Profile: Sector and Location-Specific Climate Risks and Resilience Recommendations. USAID.

Rao, T., 2021. The effectiveness of international agricultural aid in conflict-affected contexts.

Rahman, A., Westley, J., 2001. The Challenge of Ending Rural Poverty. Development Policy Review 19, 553–562. https://doi.org/10.1111/1467-7679.00152

Rhoads, E.P., Sutton, R., 2020. The (self) protection of civilians in South Sudan: Popular and community justice practices. African Affairs 119, 370–394. https://doi.org/10.1093/afraf/adaa017

Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70, 41–55. https://doi.org/10.1093/biomet/70.1.41

Rosenberg, A.M., Maluccio, J.A., Harris, J., Mwanamwenge, M., Nguyen, P.H., Tembo, G., Rawat, R., 2018. Nutrition-sensitive agricultural interventions, agricultural diversity, food access and child dietary diversity: Evidence from rural Zambia. Food Policy 80, 10–23.

Sanchez, P., Swaminathan, M.S., Dobie, P., Yuksel, N., 2005. Halving hunger - it can be done. UN Millennium Project - Task Force on Hunger.

Schreinemachers, P., Patalagsa, M.A., Islam, M.R., Uddin, M.N., Ahmad, S., Biswas, S.C., Ahmed, M.T., Yang, R.-Y., Hanson, P., Begum, S., others, 2015. The effect of women's home gardens on vegetable production and consumption in Bangladesh. Food Security 7, 97–107.

Schreinemachers, P., Patalagsa, M.A., Uddin, N., 2016. Impact and cost-effectiveness of women's training in home gardening and nutrition in Bangladesh. Journal of Development Effectiveness 8, 473–488.

Schwab, B., 2019. Comparing the Productive Effects of Cash and Food Transfers in a Crisis Setting: Evidence from a Randomised Experiment in Yemen. The Journal of Development Studies 55, 29–54. https://doi.org/10.1080/00220388.2019.1687880

Schwab, B., Margolies, A., Hoddinott, J., 2012. Impact Evaluation of Cash and Food Transfers for the Seasonal Emergency Safety Net in Hajjah and Ibb Governorates, Yemen Endline Report. International Food Policy Research Institute, Washington, DC.

Sen, A., 1981. Poverty and Famines. Oxford University Press. https://doi.org/10.1093/0198284632.001.0001

Shimeles, A., Verdier-Chouchane, A., Boly, A., 2018. Introduction: Understanding the Challenges of the Agricultural Sector in Sub-Saharan Africa, in: Shimeles, A., Verdier-Chouchane, A., Boly, A. (Eds.), Building a Resilient and Sustainable Agriculture in Sub-Saharan Africa. Springer International Publishing, Cham, pp. 1–12. https://doi.org/10.1007/978-3-319-76222-7_1

Sibhatu, K.T., Qaim, M., 2017. Rural food security, subsistence agriculture, and seasonality. PLoS ONE 12, e0186406. https://doi.org/10.1371/journal.pone.0186406

Smith, L.C., El Obeid, A.E., Jensen, H.H., 2000. The geography and causes of food insecurity in developing countries. Agricultural Economics 22, 199–215. https://doi.org/10.1111/j.1574-0862.2000.tb00018.x

Tappis, H., Doocy, S., 2018. The effectiveness and value for money of cash-based humanitarian assistance: a systematic review. Journal of Development Effectiveness 10, 121–144. https://doi.org/10.1080/19439342.2017.1363804

Tesfamariam, B.Y., Owusu-Sekyere, E., Emmanuel, D., Elizabeth, T.B., 2018. The impact of the homestead food garden programme on food security in South Africa. Food Security 10, 95–110.

Tesfaye, A., Bogale, A., Namara, R.E., Bacha, D., 2008. The impact of small-scale irrigation on household food security: The case of Filtino and Godino irrigation schemes in Ethiopia. Irrig Drainage Syst 22, 145–158. https://doi.org/10.1007/s10795-008-9047-5

Timmer, C.P., 2002. Chapter 29 Agriculture and economic development, in: Handbook of Agricultural Economics. Elsevier, pp. 1487–1546. https://doi.org/10.1016/S1574-0072(02)10011-9

UNCTAD, n.d. UN list of least developed countries | UNCTAD [WWW Document]. United Nations Conference on Trade and Development. URL https://unctad.org/topic/least-developed-countries/list (accessed 27.5.21).

UNHCR, 2021. Regional overview of the South Sudanese refugee population. United Nations High Commissioners for Refugees. URL https://data2.unhcr.org/en/dataviz/62?sv=5&geo=0

UNHCR, 2020. Regional overview of South Sudanese refugee population.

United Nations, 2020a. The Sustainable Development Goals Report.

United Nations, 2020b. List of Least Developed Countries (as of December 2020). United Nations - Commitee for Development Policy.

United Nations, 2020c. Human Development Index (HDI). URL http://hdr.undp.org/en/indicators/137506

UNMISS, 2020. Statement of the Special Representative of the UN Secretary-General David Shearer, in: Briefing to the Security Council on South Sudan.

Verwimp, P., Muñoz-Mora, J.C., 2018. Returning Home after Civil War: Food Security and Nutrition among Burundian Households. The Journal of Development Studies 54, 1019–1040. https://doi.org/10.1080/00220388.2017.1311407

Waddington, H., Snilstveit, B., Hombrados, J., Vojtkova, M., Phillips, D., Davies, P., White, H., 2014. Farmer Field Schools for Improving Farming Practices and Farmer Outcomes: A Systematic Review. Campbell Systematic Reviews 10. https://doi.org/10.4073/CSR.2014.6

Weiser, S.D., Bukusi, E.A., Steinfeld, R.L., Frongillo, E.A., Weke, E., Dworkin, S.L., Pusateri, K., Shiboski, S., Scow, K., Butler, L.M., Cohen, C.R., 2015. Shamba Maisha: randomized controlled trial of an agricultural and finance intervention to improve HIV health outcomes. AIDS 29, 1889–1894. https://doi.org/10.1097/QAD.0000000000000781

WFP, 2008. Food Consumption Analysis - Calculation and use of the food consumption score in food security analysis.

White, H., Raitzer, D.A., 2017. Impact Evaluation of Development Interventions: A Practical Guide. Asian Development Bank. https://doi.org/10.22617/TCS179188-2

Wooldridge, J., 2009. Estimating average treatment effects: difference-in-differences.

Wooldridge, J.M., 2012a. Chapter 17: Limited Dependent Variable Models and Sample Selection Corrections, in: Introductory Econometrics - A Modern Approach. pp. 583–631.

Wooldridge, J.M., 2012b. Chapter 1: The Nature of Econometrics and Economic Data, in: Introductory Econometrics - A Modern Approach. pp. 1–20.

Wooldridge, J.M., 2012c. Chapter 13: Pooling Cross Sections across Time: Simple Panel Data Methods, in: Introductory Econometrics - A Modern Approach. pp. 448–483.

Wooldridge, J.M., 2012d. Chapter 12: Serial Correlation and Heteroskedasticity in Time Series Regressions, in: Introductory Econometrics - A Modern Approach. pp. 448–483.

Wordofa, M.G., Sassi, M., 2020. Impact of agricultural interventions on food and nutrition security in Ethiopia: Uncovering pathways linking agriculture to improved nutrition. Cogent Food & Agriculture 6, 1724386. https://doi.org/10.1080/23311932.2020.1724386

World Bank, 2021a. The World Bank In South Sudan. [WWW Document]. World Bank. URL https://www.worldbank.org/en/country/southsudan/overview (accessed 4.12.2020)

World Bank, 2021b. Classification of Fragile and Conflict-Affected Situations [WWW Document]. World Bank. URL https://www.worldbank.org/en/topic/fragilityconflictviolence/brief/harmonized-list-of-fragile-situations (accessed 27.5.21).

World Bank, 2021c. South Sudan Economic Update. Fourth Edition. Pathways to Sustainable Food Security.

World Bank, 2020. South Sudan Economic Update: Peace Agreement Spurs Economic Recovery, but Poverty Remains.

World Bank, ILO, n.d. Employment in agriculture (% of total employment) (modeled ILO estimate) - South Sudan. ILOSTAT database. URL https://data.worldbank.org/indicator/SL.AGR.EMPL.ZS?locations=SS (accessed 29.1.21).

Yahaya, I., Pokharel, K.P., Alidu, A.-F., Yamoah, F.A., 2018. Sustainable agricultural intensification practices and rural food security: The case of Northwestern Ghana. BFJ 120, 468–482. https://doi.org/10.1108/BFJ-01-2017-0021

Appendix:

Tables

Table A1: Time Frame (June 2019 until the end of 2022)

	Q3/ 19	Q4/ 19	Q1/ 20	Q2/ 20	Q3/ 20	Q4/ 20	Q1/ 21	Q2/ 21	Q3 /21	Q4/ 21	Q1/ 22	Q2/ 22	Q3/ 22	Q4/ 22
Identification of beneficiaries														
Expansion of agricultural land														
Installation of demonstration gardens														
Training														
Monitoring visits														
Distribution of tools and seeds														
Installation of vegetable gar-dens														
Installation home gardens														
Cash distribution														

Notes: Only the activities relevant for this study are listed here.
Time frame is projected from Q2/21 on.

Table A2: Treatment Distribution

Kind of treatment	**Distribution key for beneficiary households**
Identification of beneficiaries	Assignment of 1,000 beneficiary farming households with special focus on vulner-able households, i.e., with (widowed) female household heads and hosts of IDP. Assignment to thirty-five farmer groups.
Expansion of agricultural land	Every household receives an additional land size between 1,050m^2 and 1,400m^2. [72]
Training	One (preferable female) household member joins two sessions of each training topic. The topics are good agricultural practices, seed selection and breeding, and improved & innovative agricultural systems
Installation of demonstration gardens	All beneficiary households are invited to visit two demonstration gardens.
Monitoring visits	A monitoring team visits the beneficiary households at least eight times a year.
Distribution of tools and seeds	Every beneficiary household receives one equipment package of tools and seeds [73].
Installation of vegetable gar-dens	The beneficiary households install gardens along the river Nile for home consump-tion and commercial purpose.
Installation home gardens	The beneficiary households renovate/install small home gardens from home con-sumption.
Cash distribution	The beneficiary households received cash according to their household size and the exposure of their neighborhood by the floods of 2020 (min. 103 EUR, max. 307 EUR[74]).

[72] Size estimation by MI.
[73] The equipment package consists of the following tools: hoe (1), fork hoe (1), slasher (1), watering can (1), rake (2); and the following seed packages: eggplant (75g), tomato (40g), Jew's melon (75g), onion (75g), amaranths (75g), okra (100g), maize (5kg), sorghum (1kg), cowpeas (1kg). Sprayer, Buckets, and deductive material are provided on farmer group level and are accessible for the beneficiary households.
[74] The exchange rate from 12.08.2020 is used (153,5) (CurrencyRate, 2021).

Table A3: Number of attritted Households by Neighborhood

	(1)	(2)
	Attritted CG	Attritted TG
Kajokeji Road	0	5
Kasire Extensions 1	7	1
Kasire Extensions 2	4	0
Kasire Nyaga 1	4	0
Kasire Nyaga 2	2	3
Kasire Nyaga 3	13	1
Kasire Nyaga 4	3	0
Kator 1	6	1
Kator 2	3	0
Kator 3	1	1
Kator 4	7	0
Logo East 1	0	0
Logo East 2	2	1
Logo East 3	2	1
Logo East 4	3	1
Logo East 5	1	3
Logo West	4	6
Lologo 1 1	3	0
Lologo 1 2	7	0
Lologo 1 3	0	0
Lologo 1 4	3	0
Lologo 2 1	5	2
Lologo 2 2	1	0
Lologo 2 3	2	0
Lologo 2 4	1	0
Lologo 2 5	7	0
Lologo 2 6	3	0
Nyori Jondoru	6	0
Tokiman East 1	6	5
Tokiman East 2	4	2
Tokiman East 3	5	2

Table A4: Baseline Group Balance of potential Covariates for PSM Subsample: Treatment Group only

	Before matching			After matching	
	No cash received	Cash received	Significance	No cash received	Significance
N	58	405		55	405
Female resp	58.6%	54.3%		55.3%	
Female HHH	51.7%	46.9%			
Age resp (years)	46.7	44.9		45.4	
Age HHH (years)	48.8	45.9		46.3	
HHH farmer	93.1%	95.3%		90.2%	
HHH literate	21.4%	34.8%	**	31.4%	
HHH married	77.6%	81,2			
HH size	8.2	8.7			
IDP/returnee HH	8.6%	6.2%		5.4%	
Land size (ha)	.48	.73	***	.66	
Land access water	51.7%	43.7%		43.7%	
Support from other organization (t=2)	24.1%	19.3%		17.8%	

Notes: Support from other organization is assessed in endline. Resp=respondent, HH=Household, HHH= Household head
Significance according to t-tests with unequal variance (welch), * p < .1, ** p < .05, *** p < .01

Table A5: Baseline Group Balance of potential Covariates for PSM Subsample: Cash receivers and Control Group

	Before matching			After matching	
	Control	Restr. TG	Significance	Control	Significance
N	384	405		384	
Female resp	74.2%	54.3	***	55.7%	
Female HHH	50.5%	46.7%			
Age resp (years)	40.2	44.9	***	46.3	
Age HHH (years)	43.5	45.9	**	46.4	
HHH farmer	85.5%	95.3%	***	94.1%	
HHH literate	32.0%	34.8%			
HHH married	80.5%	81.2%			
HH size	8.2	8.7	*	8.5	
IDP/returnee HH	9.9%	6.2%	*	6.9%	
Land size (ha)	.75	.73			
Land access water	55.5%	43.7%	***	40.4%	
Support from other organization (t=2)	6.3%	19.3%	***	15.1%	
Severe flood experience (t=2)	61.7%	65.9%		64.1%	

Notes: Support from other organization and severe flood experience are assessed in endline.
Resp=respondent, HH=Household, HHH= Household head
Significance according to t-tests with unequal variance (welch), * $p < .1$, ** $p < .05$, *** $p < .01$

Table A6: Baseline Group Balance of potential Covariates for PSM Subsample: Non-Cash receiving Beneficiaries and Control Group

	Before matching			After matching	
	Control Group	Treat-ment Group (No cash received)	Signifi-cance	Control	Signifi-cance
N	384	58		58	
Female resp	74.2%	58.6%	**	58.8%	
Female HHH	50.5%	51.3%			
Age resp (years)	40.2	46.7	***	46.8	
Age HHH (years)	43.5	48.8	***	48.6	
HHH farmer	85.2%	93.1%	**	91.1%	
HHH literate	32.0%	22.4%			
HHH married	80.5%	77.6%			
HH size	8.2	8.2			
IDP/returnee HH	9.9%	8.6%			
Land size (ha)	.75	.48	***	.53	
Land access water	55.5%	51.7%			
Support from other organi-zation (t=2)	6.3%	24.1%	***	20.8%	
Severe flood experience (t=2)	61.7%	53.4%			

Notes: Support from other organization and severe flood experience are assessed in endline.
Resp=respondent, HH=Household, HHH= Household head
Significance according to t-tests with unequal variance (welch), * p < .1, ** p < .05, *** p < .01

Table A7: Heterogenous Treatment Effects through Shocks on Food Security Outcomes

	(1) Drought	(2) Floods	(3) High agr. costs	(4) High crop pests	(5) Illness/ accident	(6) Vio-lence/ conflict	(7) Theft
Inverse Food Insecurity Experience Scale							
wave	.773***	.789***	.993***	1.004***	.862***	.703***	.756***
treat	.333**	.339	.284	.410**	.456**	.343**	.340**
wave*treat	-.006	.041	.059	-.044	-.111	-.002	-.039
	(.173)	(.270)	(.205)	(.218)	(.188)	(.180)	(.186)
shock	3.493***	-.289	.441	.524**	-.360	-.689**	-.274
shock*wave	-.958	-.135	-1.03***	-.855***	-1.034***	.047	-.481
shock*treat	1.082	.081	.329	-.083	-.602	.661	.369
shock*wave*treat	-.108	-.071	-.193	.114	.627	-.091	.364
	(.775)	(.348)	(.348)	(.340)	(.399)	(.536)	(.476)
Food Consumption Score							
wave	-1.728**	-1.636	-1.738**	-.452	-1.109	-1.171	-1.344*
treat	-1.604**	-3.546***	-1.967**	-2.178**	-1.027	-1.374*	-2.000**
wave*treat	1.156	.903	.823	-.244	.938	.882	1.317
	(.951)	(1.425)	(1.055)	(1.109)	(.999)	(.952)	(.975)
shock	4.380**	-1.207	4.509***	2.578**	2.281	3.143	-.948
shock*wave	5.136**	.436	1.332	-2.623*	-1.710	-3.142	-.218
shock*treat	3.804	3.413**	1.856	2.210	-2.360	-.649	4.154
shock*wave*treat	-3.230	.006	.277	3.318*	-.406	1.016	-2.388
	(3.327)	(1.852)	(2.092)	(1.942)	(2.457)	(3.341)	(2.758)
Minimum Dietary Diversity							
wave	-.658***	-.946***	-.823***	-.712***	-.752***	-.664***	-.683***
treat	-1.013***	-1.347***	-1.206***	-1.065***	-1.088***	-1.002***	-1.081***
wave*treat	1.076***	1.224***	1.167***	1.077***	1.291***	1.166***	1.209***
	(.146)	(.246)	(.169)	(.183)	(.157)	(.150)	(.157)
shock	.693	-.617***	-.655***	-.219	-.789***	-.439	-.515*
shock*wave	-.231	.439*	.533**	.108	.510*	-.169	.083
shock*treat	.657	.605**	.827***	.260	.748***	.445	.844**
shock*wave*treat	.879	-.158	-.131	.162	-1.116***	-.390	-.489
	(.664)	(.301)	(.313)	(.291)	(.378)	(.450)	(.345)

Notes: N=1694, Pooled OLS Regressions, SE in parentheses for relevant coefficients and clustered at household level. * p < .1, ** p < .05, *** p < .01
The variable *"shock"* is specific for each column headline.

Table A8: Baseline Balance of Neighborhoods

	Correlation with outcomes (baseline)			Group balance		
	(1) FIES	(2) FCS	(1) FIES	N (CG)	N (TG)	
Kajokeji Road	-0.121	-5.573***	-0.990***	0	59	Excluded
	(0.325)	(1.470)	(0.238)			
Kasire Extentions 1	-0.366	1.775	0.550*	26	15	
	(0.385)	(1.757)	(0.284)			
Kasire Extentions 2	1.380***	-5.210**	-0.716*	15	7	
	(0.518)	(2.366)	(0.384)			
Kasire Nyaga 1	-2.852***	-2.984	-0.676	4	10	
	(0.642)	(2.958)	(0.479)			
Kasire Nyaga 2	1.250***	-1.217	-1.145***	1	25	
	(0.478)	(2.187)	(0.352)			
Kasire Nyaga 3	0.611	-3.450	0.909***	22	5	
	(0.471)	(2.145)	(0.347)			
Kasire Nyaga 4	0.737	1.003	-0.383	8	0	Excluded
	(0.855)	(3.901)	(0.632)			
Kator 1	-0.716	5.849***	1.214***	20	9	
	(0.454)	(2.065)	(0.334)			
Kator 2	-2.307***	1.522	-0.066	10	28	
	(0.392)	(1.822)	(0.295)			
Kator 3	0.371	-3.428	-0.514	1	24	
	(0.489)	(2.226)	(0.361)			
Kator 4	-1.028***	-0.278	-0.121	13	36	
	(0.353)	(1.616)	(0.262)			
Logo East 1	1.296*	-5.066	-0.788	1	9	
	(0.765)	(3.489)	(0.566)			
Logo East 2	0.141	-0.899	-0.0354	17	12	
	(0.455)	(2.075)	(0.336)			
Logo East 3	0.236	3.476	-0.337	6	14	
	(0.545)	(2.482)	(0.403)			
Logo East 4	-1.731***	2.281	1.299***	6	19	
	(0.485)	(2.228)	(0.359)			
Logo East 5	0.437	1.507	0.110	13	20	
	(0.427)	(1.949)	(0.316)			
Logo West	-0.150	-2.126	-0.814***	15	43	
	(0.328)	(1.492)	(0.241)			
Lologo 1 1	0.775	-1.057	0.521	16	2	
	(0.573)	(2.616)	(0.424)			
Lologo 1 2	1.193**	1.557	0.404	22	0	Excluded
	(0.519)	(2.372)	(0.384)			
Lologo 1 3	1.286**	3.229	-0.160	4	14	
	(0.572)	(2.614)	(0.424)			
Lologo 1 4	0.165	-3.811*	1.159***	27	1	
	(0.463)	(2.106)	(0.340)			
Lologo 2 1	0.439	-1.524	-0.119	15	19	
	(0.421)	(1.922)	(0.312)			

Lologo 2 2	-0.668	-4.297*	0.102	13	10	
	(0.509)	(2.317)	(0.376)			
Lologo 2 3	0.557	-2.823	0.564	15	0	Excluded
	(0.627)	(2.859)	(0.463)			
Lologo 2 4	-1.766***	4.695**	0.720**	26	0	Excluded
	(0.476)	(2.182)	(0.354)			
Lologo 2 5	0.964	6.746**	-0.589	15	0	Excluded
	(0.627)	(2.852)	(0.463)			
Lologo 2 6	1.384*	0.723	0.965	9	0	Excluded
	(0.806)	(3.680)	(0.596)			
Nyori Jondoru	0.710	5.665	-0.046	9	0	Excluded
	(0.807)	(3.675)	(0.597)			
Tokiman East 1	0.872**	3.160**	-0.530**	5	45	
	(0.350)	(1.597)	(0.259)			
Tokiman East 2	0.741**	3.112**	0.598**	23	28	
	(0.347)	(1.583)	(0.256)			
Tokiman East 3	-0.275	2.860	0.123	7	9	
	(0.608)	(2.770)	(0.449)			
N	847	847	847	384	463	

Notes. Standard errors in parentheses, * p < .1, ** p < .05, *** p < .01, Excluded= Neighborhoods are excluded from the analysis because of no representation of one group.

Figures

Figure A1: Household Flood Exposure
Snapshot from own visualization with QGIS and OpenStreetMap
Information provided by Malteser International
2021

Figure A2: Clustering Approach on Neighborhood Level
Snapshot from own visualization with QGIS and OpenStreetMap
Left: zoomed & unlabeled, right: labelled

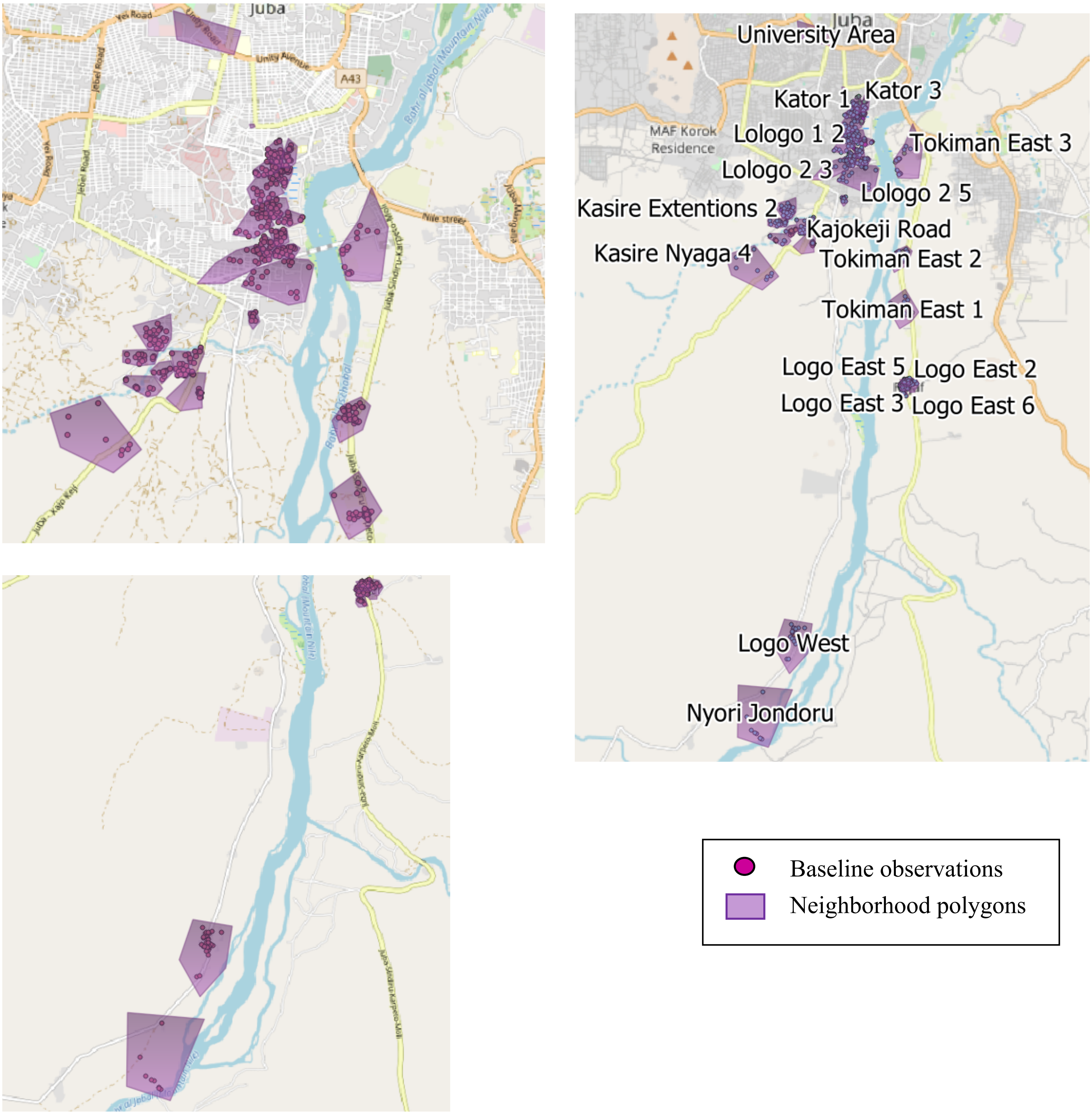


Figure A3: Distribution of Propensity Score across Subgroups:

Treatment Group that received Cash vs. Treatment Group that did not

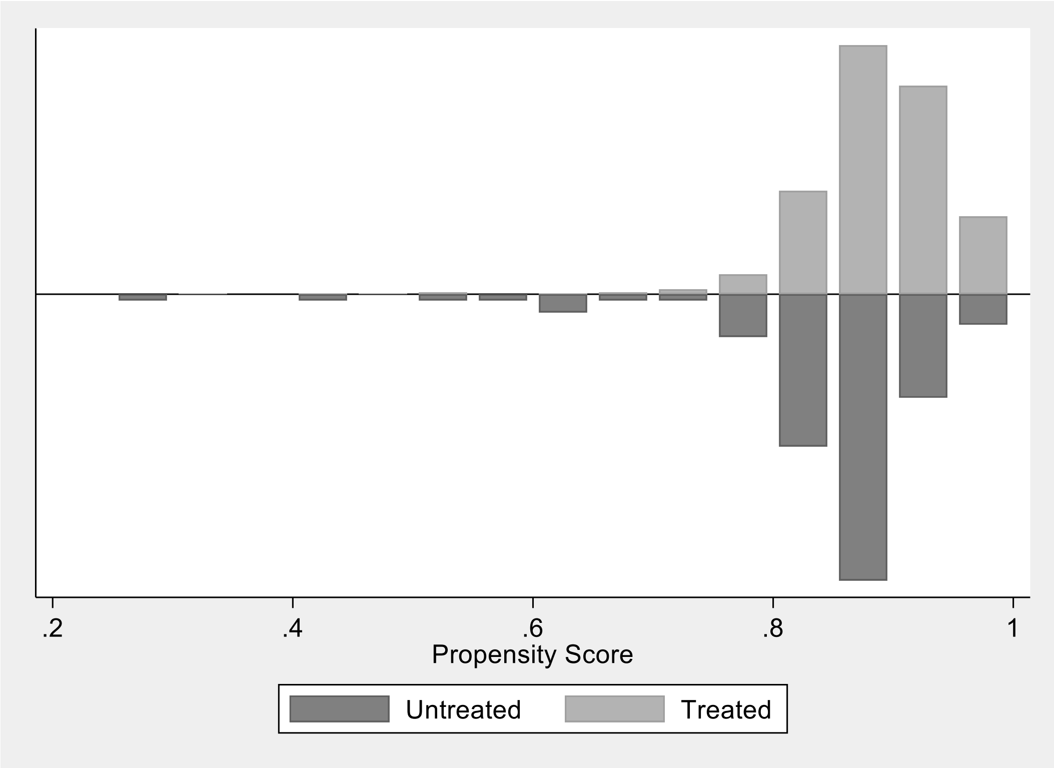


Figure A4: Distribution of Propensity Score across Subgroups:

Treatment Group that received Cash vs. Control Group

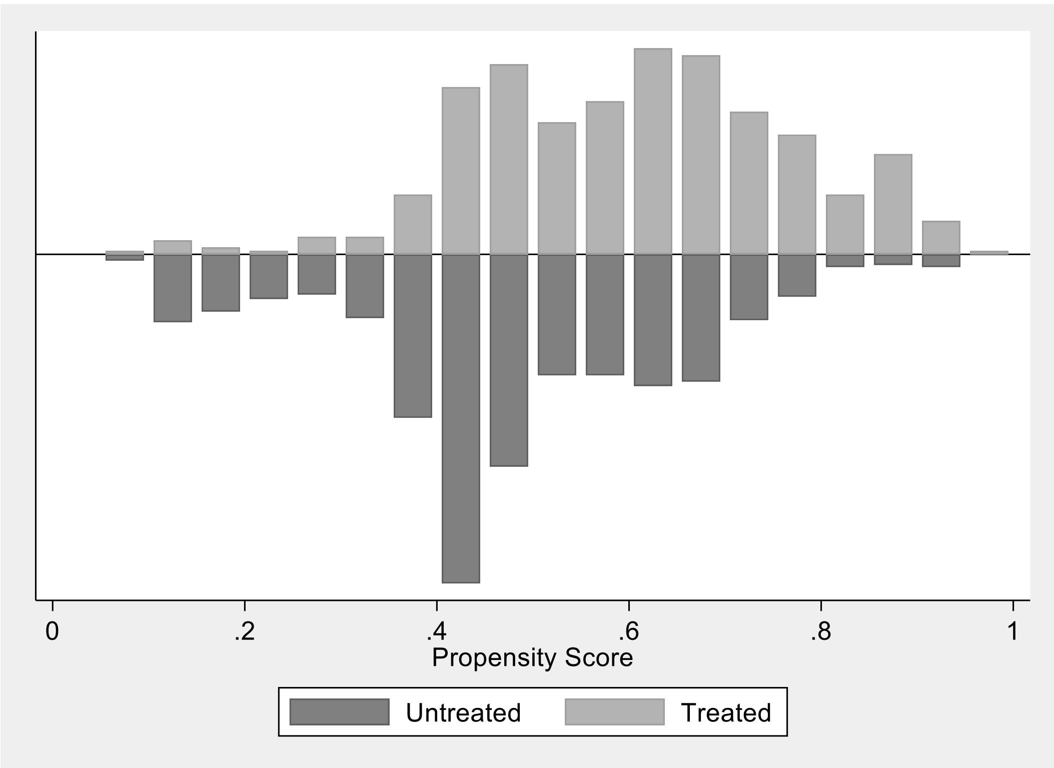


Figure A5: Distribution of Propensity Score across Subgroups:

Treatment Group that did not receive Cash vs. Control Group

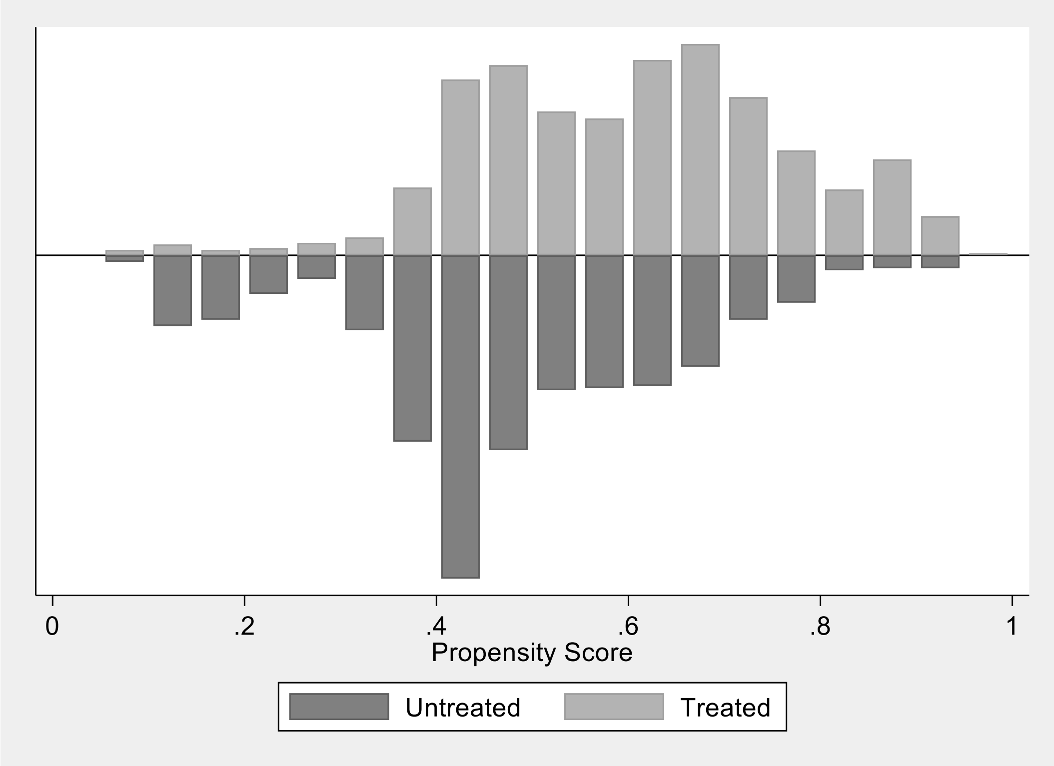


Figure A6: Distribution of Propensity Score across Groups:

Inclusion of Neighborhoods