Assessing the Population-level Impact of USAID's Feed the Future Initiative: Insights from Bangladesh Amidst Global Aid Reductions*

Akhter U. Ahmed[a]

M. Mehrab Bakhtiar[a]

Daniel O. Gilligan[a]

Giang Thai[b]

Abstract:
International development assistance provides substantial investment in LMICs and is provided by donor countries for political, economic and humanitarian reasons. Most development assistance is earmarked for specific sectors or ministries, its delivery often managed by international NGOs or other proxies. Evidence on the effectiveness of this assistance focuses on evaluations of individual projects or modeling of return on foreign assistance, but well-identified evidence on the impact of an entire program of bilateral development assistance is rare. Using a quasi-experimental matching design, we provide the first evidence of the impact of USAID's largest development assistance program in one of its highest priority countries: the Feed the Future program in Bangladesh. We examine the population-level impact of FTF investments on people living in its 20 target districts on food security, poverty, child malnutrition, agricultural production and women's empowerment in Bangladesh from 2011-2019. We find plausible positive impacts on food security, agricultural production and empowerment, impacts that largely reflect USAID investment priorities. We find no effect on prevalence of poverty or child malnutrition. Results suggest that impacts on these outcomes would require larger sustained investments as well as more targeted assistance to the poor.

[a]Poverty, Health and Nutrition Division, International Food Policy Research Institute (IFPRI), Washington DC, United States
[b]Department of Applied Economics, University of Minnesota, St. Paul, MN

Keywords: Feed the Future, international development assistance, impact of aid, Bangladesh
JEL Codes: O19, O22, F35, Q18

* This work was undertaken as part of the CGIAR Research Program on Policies, Institutions, and Markets (PIM) led by the International Food Policy Research Institute (IFPRI). We gratefully acknowledge funding from the United States Agency for International Development (USAID) through agreement number AID-BFS-IO-14-00002 and from PIM. We have received helpful advice and comments from Anne Swindale, Chris Hillbruner, Lesley Perlman, and an anonymous reviewer. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of USAID, IFPRI or the CGIAR. Corresponding author: Daniel O. Gilligan, d.gilligan@cgiar.org, International Food Policy Research Institute (IFPRI), 1201 Eye St, Washington DC 20006-1002, United States.

1 Introduction

International development assistance, a cornerstone in the global economic landscape, plays a pivotal role in fostering economic growth, reducing poverty, and improving nutrition in low- and middle-income countries (LMICs). The allocation of over $150 billion annually in international assistance underscores its significance (OECD, 2019). This substantial investment reflects a global commitment to addressing the multifaceted challenges of poverty and underdevelopment. While primarily aimed at bolstering economic and political stability in recipient countries, development assistance also serves strategic interests for donor nations, creating new markets and strengthening economic ties (Dreher et al., 2017; Berthélemy, 2006; Arndt et al., 2015).

The effectiveness of development assistance, however, remains a subject of intense debate. Donor nations are under increasing pressure to justify these investments, necessitating that funding is used for programs that have demonstrable impact (Dietrich, 2013; Easterly, 2006). Critics argue that aid can perpetuate dependency, foster corruption, and fail to address the root causes of poverty (Moyo, 2009; Collier, 2007). Conversely, proponents highlight the potential for aid to catalyze development, particularly when aligned with effective policies and governance (Sachs, 2005; Clemens et al., 2012). While numerous studies have evaluated the impacts of specific projects (Riddell, 2007; Burnside & Dollar, 2000), comprehensive assessments of large-scale programs, such as the United States Agency for International Development's (USAID) Feed the Future (FTF) program, at the country level are scarce.

FTF has garnered significant attention as USAID's flagship initiative aimed at tackling global hunger and food insecurity since it was launched in 2010. As of 2023, the program has invested over $11 billion across 19 focus countries, and has also supported projects in aligned and regional programs (USAID, 2023). The stated objectives of the program are to reduce poverty and child stunting at the population level in its targeted operating area, known as the Zone of Influence (ZOI), and to increase agricultural productivity and diversification, promote food security, and increase women's empowerment. FTF attempts to achieve these goals in its focus countries through a web of projects and activities, such as Bangladesh Nutrition Activity, Bangladesh Rice and Diversified Crops, Livestock for Improved Nutrition, etc. Rigorous evaluation of FTF would provide vital evidence about whether these investments are having a meaningful impact on its main objectives, and would inform whether further investment is justified, but so far no rigorous impact evaluation of an FTF program at the country level exists (Bazzi & Clemens, 2013; Chandy & Kharas, 2011).

This paper seeks to fill this evidence gap by reporting on the first rigorous impact evaluation of a USAID FTF program at the country level, in this case the program in Bangladesh. Bangladesh is the largest recipient of Feed the Future investment, receiving over $3.3 billion in funding since 2010 (USAID, 2022). Employing panel data of nationally- and regionally-representative rural households from the Bangladesh Integrated Household Survey (BIHS), this study utilizes a quasi-experimental design to assess FTF's impact within its Zone of Influence (ZOI) at the population level. This methodological approach allows for a more nuanced understanding of the program's effects, accounting for both direct and indirect impacts on targeted populations (Khandker et al., 2010; Gertler et al., 2016). We compare outcomes for households within the FTF ZOI against a counterfactual of households outside this zone, focusing on poverty, hunger, food insecurity, child stunting, women's empowerment, technological

adoption, agricultural yield, and child feeding behaviors. The comprehensive nature of this analysis provides valuable insights into the multifaceted impacts of development assistance, contributing to a more informed and nuanced understanding of its effectiveness in LMICs (De Janvry & Sadoulet, 2010; Barrett et al., 2012).

We find that the FTF program in Bangladesh had significant positive impacts on agricultural diversification, food security, and women's empowerment from 2011-2019 but no effects on poverty reduction and child malnutrition. For food security, households within ZOI were 8.5 percentage points less likely to experience any food insecurity and 9.4 percentage points less likely to experience mild food insecurity compared to those living outside of ZOI. With respect to women's empowerment, we find that both men and women in ZOI households were more empowered. In particular, we see a statistically significant increase in the probability that a woman is empowered by 5.1 percent and the probability that men are empowered by a statistically significant 5.7 percent. Estimated impacts on the gap between a man's and woman's 5DE score and on the probability that the household reached gender parity show small positive effects.

For agriculture, we find that Feed the Future led to a reduction in the share of rice in total planted area of 12.2 percentage points and a reduction in the share of high-yielding variety (HYV) rice in total planted area of 18.6 percentage points. The former effect may reflect the impact of Feed the Future on crop choice. Indeed, Feed the Future's efforts to promote crop diversification led to a large, significant impact on the number of crops grown, by 0.5 crops per farmer. We also see that Feed the Future led to large and significant increases in rice yields for TAman and Boro HYV varieties. These results for agricultural production show that Feed the Future played a key role in improving yields of HYV rice varieties and promoting greater crop diversity, during a period in which yields were falling overall.

The program's influence on agricultural practices has been particularly noteworthy. It has led to a shift in crop patterns, with a marked increase in the cultivation of high-yielding and diverse crop varieties. This shift not only contributes to food security but also enhances the resilience of agricultural systems against market fluctuations and environmental stresses (Minten et al., 2013; Spielman et al., 2017). However, the impact on poverty reduction and child stunting was less pronounced, suggesting the need for more targeted interventions in these areas (Headey et al., 2012; Hoddinott et al., 2013). While the program has made strides in improving overall food availability, its effects on the most vulnerable populations, particularly in terms of nutritional outcomes for children, have been limited. This aligns with findings from other studies indicating that agricultural growth does not automatically translate into proportional improvements in nutrition (Ruel et al., 2018; Webb & Block, 2012).

This research contributes to the literature by providing empirically robust evidence of the causal impact of a country-level Overseas Development Assistance (ODA) program. It elucidates potential pathways through which agricultural and empowerment interventions can influence broader socio-economic outcomes. Our analysis reveals that interventions aimed at women's empowerment have led to increased participation in agricultural decision-making and labor markets, echoing the findings of Quisumbing et al. (2015) and Peterman et al. (2014) on the positive ripple effects of women's empowerment in rural settings. This contributes to the debate on the effectiveness of sector-specific versus holistic development approaches (Dorosh & Thurlow, 2014; Barrett et al., 2010). The findings offer practical insights for governments and international donors in optimizing the impact of ODA. They

underscore the importance of integrating agricultural development with specific nutrition and poverty alleviation strategies to achieve more comprehensive development outcomes (De Janvry & Sadoulet, 2010; Malapit et al., 2015).

While the FTF program in Bangladesh has made notable strides in certain areas, the lack of impact on poverty and nutrition highlight the complexities of development assistance. These findings underscore the importance of continuous evaluation and adaptation of development programs to ensure their effectiveness and relevance in achieving their intended outcomes. These results also emphasize the importance for donor countries to set realistic, achievable objectives for development assistance. The failure of FTF in Bangladesh to reduce poverty and child stunting at the population level has undermined support for the program, despite its important impacts on agriculture, food security and women's empowerment. Targeting reductions in poverty and malnutrition in the targeted households is more realistic; evidence of such effects would likely provide further support for the program. The mixed impacts in our study suggest that future programs should consider more integrated approaches, targeted more directly at poor households and in areas with more child malnutrition, that combine agricultural development with targeted nutrition and social protection interventions (Bhutta et al., 2013; Leroy et al., 2016). Such integrated strategies are crucial for addressing the multifaceted challenges of poverty and undernutrition in LMICs.

The rest of this paper is organized as follows: Section 2 provides a broad overview of the Feed the Future program and Section 3 introduces the BIHS data. Section 4 outlines the empirical approach and Section 5 presents and discusses the impact results. Section 6 concludes.

2 Feed the Future in Bangladesh

2.1 The Feed the Future ZOI

The Bangladesh Feed the Future ZOI includes rural areas of 20 districts consisting of 120 upazilas (sub-districts) in three divisions in the south and southwest region of the country:

1. Barishal Division (six districts): Barguna, Barishal, Bhola, Jhalakati, Patuakhali, and Pirojpur;
2. Dhaka Division (five districts): Faridpur, Gopalganj, Madaripur, Rajbari, and Shariatpur; and
3. Khulna Division (nine districts): Bagerhat, Chuadanga, Jessore, Jhenaidah, Khulna, Magura, Meherpur, Narail, and Satkhira.

Feed the Future targets investments in the ZOI that have the greatest potential to increase productivity of rice production and those that promote agricultural diversification toward high-value agriculture. Feed the Future also includes objectives to improve maternal and child nutritional status and to strengthen women's empowerment. Geographically, the Feed the Future ZOI is situated in the South of Bangladesh, an area characterized by high levels of poverty and malnutrition. More than 28 million people live in Feed the Future's target regions in Bangladesh, where food security and nutrition face considerable challenges associated with scarce water resources, a rising sea level, vulnerability to extreme shocks, and changing weather patterns. A map of the Feed the Future ZOI in Bangladesh is provided in the online appendix (Figure 1).

2.2 Feed the Future activities

Feed the Future delivered many projects during the period between the BIHS 2011/12 and 2018/19 survey rounds, which covers most of Feed the Future Phase One and the beginning of Phase Two. These projects included the following:

1. Bangladesh Nutrition Activity
2. Bangladesh Rice and Diversified Crops
3. Livestock for Improved Nutrition
4. Social and Behavior change communication Activity (GH activity)
5. Nobo Jarta (Food for Peace)
6. Strengthening Household Ability to Respond to Development Opportunities (SHOUHARDO) III (Food for Peace)
7. Sustainable Agriculture and Production Linked to Improved Nutrition Status, Resilience, and Gender Equity (SAPLING) (Food for Peace)
8. Policy Research and Strategy Support Program (PRSSP) - Phase II
9. Bangladesh Improving Trade and Business Enabling Environment Activity
10. Strengthening Multi-Sectoral Nutrition Programming through Implementation Science (GH activity)
11. Bangladesh South Asia Eggplant Improvement Partnership
12. Bangladesh Biotechnology Partnership
13. Bangladesh Aquaculture and Nutrition Activity
14. Chittagong Hill Tracts Watershed Co-Management Activity

Although some of these projects were formally developed under Food for Peace, they are considered part of the Feed the Future portfolio when operated in the ZOI. More information about Feed the Future programming in Bangladesh and a current list of programs are available in the country Feed the Future webpage.[2] Numerous reports and other studies (too many to list here) have documented Feed the Future activities in Bangladesh, the status of agriculture and food security in the ZOI and many performance indicators. In addition, some activities within Feed the Future have been studied in detail, such as the Agriculture Value Chain (AVC) component (IFPRI 2020). However, as with Feed the Future activities in other countries, there has not yet been an evaluation of the impact of the comprehensive set of Feed the Future activities.

3. The Bangladesh Integrated Household Survey (BIHS) Data and Sample

3.1 BIHS Data

BIHS is a panel survey that collected data on the same households over time. IFPRI administered three panel survey rounds: 2011/2012 (baseline), 2015 (midline), and 2018/2019 (endline). The BIHS data used in this analysis were collected in the baseline and endline survey rounds. The Feed the Future ZOI was designed as a separate sample stratum in the BIHS. The BIHS 2011/2012 baseline sample includes stratification that makes the survey statistically representative at the following levels: (1) nationally representative of rural Bangladesh; (2) representative of rural areas of each of the seven administrative divisions of the country: Barishal, Chittagong, Dhaka, Khulna, Rajshahi, Rangpur, and Sylhet; and (3)

[2] https://www.feedthefuture.gov/country/bangladesh/

representative of the Feed the Future ZOI.[3] USAID provided the list of ZOI locations (districts and upazilas). Using this list, a statistician sampled the ZOI stratum separately for its statistical representativeness.[4] The fieldwork for the national level and ZOI-level surveys were carried out as part of the same survey exercise, though there were some differences in the timing of the fieldwork for each component. The ZOI baseline took place in October-November 2011, while the baseline survey for the sample strata outside the ZOI was conducted from December 2011 – March 2012. The ZOI endline was conducted in November 2018 – February 2019, and the endline BIHS interviews outside the Feed the Future ZOI were conducted in January-April 2019.

The baseline BIHS sample includes 6,500 households in 325 primary sampling units (PSUs, villages). The sample design of the BIHS followed a stratified random sampling in two stages—selection of PSUs and selection of households within each PSU—using the sampling frame developed from the community series of the 2001 population census of Bangladesh. Later, sampling weights were adjusted based on the latest population census of 2011. The domain of the national survey was the rural areas of the entire country, and the domain of the Feed the Future ZOI was the rural areas of 120 upazilas in 20 districts.

In the first stage of sampling, the total BIHS sample of 325 PSUs were allocated among the eight strata (seven divisions and the ZOI) with probability proportional to size (size being the number of households in each stratum), which resulted in the following distribution: 21 PSUs in Barishal, 48 in Chittagong, 87 in Dhaka, 27 in Khulna, 29 in Rajshahi, 27 in Rangpur, 36 in Sylhet, and 50 in the Feed the Future ZOI. In the second stage, 20 households were randomly selected from each PSU.

3.2 The Feed the Future ZOI BIHS sample

The ZOI stratum is situated in three of the seven divisions (strata): Barishal, Dhaka, and Khulna. There are 102 PSUs in the ZOI sample. Initially, the ZOI stratum had a sample of 1,000 households in 50 PSUs: 14 PSUs in Barishal, 12 PSUs in Dhaka, and 24 PSUs in Khulna. However, IFPRI researchers noticed that the sample size was inadequate for certain disaggregated analyses of the data. To obtain more robust estimates of disaggregated analysis, the researchers expanded the ZOI sample of households by adding 52 PSUs (with 1,040 sample households) from the nationally representative BIHS sample. These additional PSUs are situated in the ZOI as follows: in Barishal (17 PSUs), Dhaka (13 PSUs), and Khulna (22 PSUs). Since the sampling frame of the BIHS has the ZOI stratum and the seven strata represent the seven divisions, use of the additional BIHS sample from the three divisional strata required estimation of appropriate sampling weights to obtain results that are statistically representative of the ZOI.

In the end, the 2011/2012 baseline sample in the ZOI includes 2,040 households (1,000 households in the original ZOI sample and 1,040 additional sample households) in 102 PSUs. The ZOI endline survey in 2018/2019 sought to reinterview all households from the 2011/2012 ZOI baseline survey, making a panel dataset. At endline, there are a total of 2,064 interviews in Feed the Future ZOI households, which includes 1,908 households from the baseline and 156 additional households that split from the original baseline sample. Based on these figures, the attrition rate for households in the baseline Feed the Future

[3] The administrative structure of Bangladesh consists of divisions, districts, upazilas (sub-districts), and unions, in decreasing order of size. There are seven divisions, 64 districts, 484 upazilas, and 4,498 unions.

[4] The BIHS sampling was done by the former Chief Statistician at the Bangladesh Bureau of Statistics, Ministry of Planning.

ZOI sample was 6.5 percent by endline (93.5 percent of baseline Feed the Future ZOI households were interviewed at endline). By far the most common reason for attrition was migration.[5]

3.3 The nationally representative BIHS sample

The impact analysis uses the full BIHS sample, which included 6,503 households at baseline. This allows the Feed the Future treatment sample in the impact analysis to be based on a representative sample of the Feed the Future ZOI, and the comparison sample is drawn from households across Bangladesh outside the Feed the Future ZOI. In the full sample, there are 102 PSUs in the ZOI stratum and 223 PSUs in the rest of the BIHS sample. The 102 PSUs in the ZOI include 50 PSUs in the original Feed the Future sample, plus 52 PSUs that were part of the nationally representative sample. The full BIHS sample includes 5,503 households in 275 PSUs in the nationally representative sample plus 1,000 households in the original 50 ZOI PSUs. Attrition in the sample is low. At endline, the survey team interviewed 5,111 households from the 5,503 households in the baseline nationally representative sample.[6] This indicates an attrition rate of 7.1 percent from baseline to endline.[7] In order to increase the statistical power of the impact analysis, we also include the 1,000 baseline households in 50 PSUs from the original Feed the Future sample.

3.4 Trends in outcomes over time

In order to better understand the context for the impact analysis, it is useful to summarize the trends in outcomes from 2011/12 to 2018/19 for the ZOI and non-ZOI samples. These trends are reported for selected outcomes in Appendix Table B.1, separately for households in the ZOI sample and for those outside the ZOI sample.

Appendix Table B.1 reports two measures of statistical significance for the hypothesis test that the mean outcomes are equal over time, from 2011/12 to 2018/19. P-values are from a t-test of equality of means. A p-value below 0.05 indicates a significant difference in means at the 5 percent level. The normalized difference for each outcome in each sample is defined as the difference in sample means over time (mean outcome in 2018/19 minus the mean outcome in 2011/12) divided by the average standard deviation in 2011/12 and 2018/19 (Imbens 2015). The normalized difference is a measure of significance that is not affected by sample size. As suggested by Imbens and Rubin (2015), we interpret normalized differences of 0.25 or above as indicating a significant difference in means between the baseline and endline surveys.

Appendix Table B.1 shows that at the time of the baseline survey, the Feed the Future ZOI was similar in poverty prevalence to the regions outside the ZOI. Poverty prevalence (per capita expenditure per day below $1.25 (2005 PPP)) was 37.4 percent in the ZOI, and was slightly lower, at 35.1 percent, outside the ZOI.[8] By endline in 2018/19, poverty had fallen substantially both in the ZOI and outside the ZOI, to 21.5

[5] We present a detailed analysis of attrition in Section 3.7.
[6] The endline survey also included 493 additional households that split from the baseline sample households, mostly due to marriage.
[7] We test for selective attrition in the full sample used in the impact analysis. However, by matching on observables, the matching model is unlikely to produce biased estimates of impact on the remaining, non-attrited sample. Still, it is important to consider if those attriting (mostly due to migration) are systematically different than those that remain in the sample.
[8] This analysis uses the $1.25 (2005 PPP) cutoff for per capita expenditure per person per day as the poverty line, which was the primary outcome indicator for poverty in the first phase of Feed the Future.

percent and 19.4 percent, respectively. The decline in poverty in the ZOI during this period was statistically significant as determined by both the p-value (p=0.000) and the normalized difference (nd=0.354). Similarly, in the sample outside the ZOI, the decline in poverty was statistically significant as determined by both the p-value (p=0.000) and the normalized difference (nd=0.359).

The data show that the Feed the Future ZOI included lower-performing regions in agriculture at the start of Feed the Future. At baseline, the proportion of farmers cultivating high-yielding variety (HYV) rice was somewhat higher outside the ZOI (86.6 percent) than inside the ZOI (80.0 percent) which reflects greater investment in more productive rice varieties in the somewhat better off regions outside the ZOI. In addition, HYV adoption is somewhat lower in the ZOI than outside the ZOI because the coastal regions in the ZOI are less well suited to the HYV rice varieties, particularly for Boro rice, due to higher salinity of soils near the coastline. Some local rice varieties have been bred to better withstand this salinity and so are preferred by some farmers near the coast. The proportion of farmers growing HYV rice varieties remained relatively stable through the endline, at 80.7 percent and 88.3 percent in the ZOI and outside the ZOI, respectively.

In the ZOI sample, the share of households growing TAus rice fell from 12.0 percent to 8.3 percent, while the share of households growing TAman rice grew sharply from 39.3 percent to 56.2 percent.[9] This increase in TAman cultivation was significant (p=0,000; nd=-0.344). The share of households growing boro rice in the ZOI fell from 55.6 percent to 49.2 percent, but this decline was not significant according to the normalized difference. Outside the ZOI, the trend in rice cultivation was similar, with the share of households growing TAus rice falling from 10 percent to 5.9 percent, the share of households growing TAman rice rising from 51.6 percent to 61.7 percent, and the share of households growing boro rice falling from 73.0 percent to 70.4 percent. These changes are statistically significant according to the t-tests, but not according to the normalized differences. As expected, rice is the dominant crop in both samples: the share of rice in total crop area planted in the ZOI was 66.5 percent at baseline and rose slightly to 68.0 percent, a change that was not significant. Outside the ZOI, the share of crop area planted in rice was unchanged between baseline and endline at 79.6 percent.

At baseline, rice yields on average were similar in the ZOI to outside the ZOI (at 2.83 and 2.80 metric tons/ha, respectively). Yields increased in both samples by endline, rising to 3.08 tons/ha in the ZOI (not significant according to the nd) and 3.13 tons/ha outside the ZOI (a significant change according to the nd). For the HYV rice crops, TAus, TAman and Boro, baseline yields were higher in the ZOI than outside the ZOI, particularly so for TAus and TAman rice. Average yields for hybrid and HYV varieties (excluding local varieties) were modestly higher at baseline in the ZOI than outside the ZOI. By the 2018/19 survey round, HYV rice yields had increased both in the ZOI and outside the ZOI, with average HYV yields rising 7.4 percent in the ZOI and 7.5 percent outside the ZOI on average.

[9] There are three rice-growing seasons in Bangladesh: aus, aman, and boro. Aus is the pre-monsoon rice-growing season under rainfed conditions. Aus rice (T. Aus when transplanted) is sown March-April and harvested between July and August. The monsoon-season rainfed rice is aman, which is planted in two ways: direct-seeding (broadcast or B. Aman) with aus in March-April and transplanted (T. Aman) between July and August. Both types are harvested from November through December. Boro is the dry-season irrigated rice planted from December to early February and harvested between April and June. In 2018/2019, the total production of rice in Bangladesh was 36.4 million metric tons (MT) of milled rice, of which boro rice accounted for 54 percent; aman rice, 38 percent; and aus rice, 8 percent.

Most farmers in Bangladesh only grow a couple of crops, reflecting the overwhelming importance of rice in farming. Table 1 shows that, at baseline in 2011/12, crop diversification was somewhat higher in the ZOI than outside the ZOI: farmers grew 2.3 crops on average in the ZOI and 1.9 crops on average outside the ZOI. By endline in 2018/19, crop choice had become even more concentrated, with 2.1 crops grown in the ZOI and 1.7 crops grown outside the ZOI on average.

We have seen that the ZOI faced some disadvantages in poverty prevalence and average agricultural productivity but also some advantages in terms of HYV rice yields and crop diversification compared to elsewhere in Bangladesh. Against this complex backdrop, child nutrition outcomes were somewhat better at baseline in the ZOI than outside the ZOI. Appendix Table B.1 shows higher mean height-for-age z-score (HAZ) for children under 5 in the ZOI (-1.78) than outside the ZOI (-1.85), and lower stunting prevalence in the ZOI (43 percent) than outside the ZOI (48 percent). Similarly, for children under age 2, stunting prevalence was lower at baseline in the ZOI (33.1 percent) than outside the ZOI (38.3 percent). By endline, child stunting had fallen sharply across Bangladesh. For children under 5, stunting had fallen to 30.1 percent in the ZOI and 32.9 percent outside the ZOI; both of these declines were statistically significant.

These trends show substantial improvements in wellbeing across Bangladesh during this period, with both poverty and child malnutrition falling sharply. For Feed the Future to have had a positive impact, the gains in the ZOI would have needed to be even larger than outside the ZOI for comparable households. The trends here show that hunger fell more sharply in the Feed the Future ZOI than elsewhere while gains in other outcomes were similar both inside and outside the ZOI. However, the impact evaluation will derive estimates about whether the gains in outcomes are greater for households in the ZOI than for similar households living elsewhere in Bangladesh.

3.5 Outcome variables in 2018/19

We now summarize the full set of outcome variables at endline that will be used in the impact analysis, for ZOI and non-ZOI regions. The outcome variables are based on the list of outcomes agreed to with USAID, as listed in the research protocol (Ahmed et al. 2020). For the impact analysis, we grouped these variables into three outcome groups: household welfare; child nutrition and feeding practices; and agricultural production. The outcomes in each group are related and are considered to have similar underlying determinants, so they can be included in the same family of outcomes for the impact analysis. These groupings are made for practical reasons, so that the same set of covariates is used within each family of outcomes. While it would be feasible to have different covariates in the matching model for each outcome, this would raise concerns that the differences in covariates between models for similar outcomes are arbitrary. A summary of the outcome types in each outcome groups is as follows:

Outcome group 1: household welfare: poverty, hunger, food security & women's empowerment	• Poverty prevalence: daily per capita expenditure per person < $1.25 PPP[10] • Household Hunger Scale • Food Insecurity Experience Scale • Women's empowerment (Abbreviated Women's Empowerment in Agriculture Index, A-WEAI)

[10] PPP, purchasing power parity, is a metric used to compare the value of equivalent goods across countries, adjusting for differences in prices.

Outcome group 2: child nutrition and feeding practices	• mean HAZ and stunting prevalence for children age 0-59 months and 0-23 months • Infant and young child feeding (IYCF) practices
Outcome group 3: agricultural production	• adoption of high yielding variety (HYV) rice (any HYV adoption and specific varieties) • average rice yield per hectare in the previous year • crop diversity

See Appendix Table A.1 for the complete list of outcome variables. We measure food insecurity using the FAO's Food Insecurity Experience Scale (FIES), which measures food insecurity as a count variable based on the number of food insecurity measures experienced by the household in the last 12 months.[11] The FIES is based on Yes/No responses to the following eight yes or no questions.

During the last 12 months, was there a time when, because of lack of money or other resources:

You were worried you would not have enough food to eat?

You were unable to eat healthy and nutritious food?

You ate only a few kinds of foods?

You had to skip a meal?

You ate less than you thought you should?

Your household ran out of food?

You were hungry but did not eat?

You went without eating for a whole day?

Next, we turn to the measures of women's empowerment, which were taken from the Abbreviated Women's Empowerment in Agriculture Index (A-WEAI) (see Malapit et al. 2017). The A-WEAI includes measures of women's empowerment across five domains: production, resources, income, leadership and time. In addition, we report A-WEAI composite indicators, including the five domain empowerment (5DE) score, which is a weighted average of scores across five domains of empowerment including production, resources, income, leadership, and time use. The 5DE score ranges from 0 to 1. A person is rated as empowered if their 5DE score is 0.80 or higher. A household is considered to have reached gender parity if the woman's 5DE score is at least as high as that of her partner, or if the woman is already empowered in the 5DE score. There is no distinctive pattern between the ZOI and non-ZOI samples for women or for men in the A-WEAI indicators.

3.6 Contextual variables used in matching, from 2011/12

The baseline contextual variables used in the matching models are summarized in Table 1. These variables include measures of household demographics, household head occupation, assets, housing characteristics, exposure to shocks, village characteristics (measured using the means of the household survey data by village), and maternal and child characteristics. Household heads in the ZOI are older, more likely to be male, are more likely to work in agriculture, and have higher education on average than outside the ZOI (these differences are statistically significant based on the p-value). Several of the asset and wealth measures are higher in the ZOI than outside the ZOI at baseline, including for area of land

[11] http://www.fao.org/policy-support/tools-and-publications/resources-details/en/c/1236494/

owned, livestock owned in Tropical Livestock Units (TLUs) and a household wealth index. Households in the ZOI are also significantly more likely to report a negative shock in the past 5 years (p-value<0.05).

3.7 Tests for selective attrition

Although attrition from the BIHS sample between baseline and endline is low, at 6.5 percent in the Feed the Future ZOI sample and 7.1 percent in the rural nationally representative sample over 8 years, we examine how attrited households differ from the analysis sample of households that remained in the panel until endline. Table 2 reports summary statistics by attrited status for the complete set of control variables for the ZOI and non-ZOI samples.

Overall, we find moderate levels of selection into attrition. Female headed households were significantly more likely to attrit in both the ZOI and non-ZOI samples. Smaller households were also more likely to attrit. Outside the ZOI, households whose head's primary occupation was in agriculture were more likely to remain in the sample, likely due to their ties to their land. Wage workers are more likely to attrit in the non-ZOI sample and less likely to attrit in the ZOI sample. Also, households with less land and fewer assets are more likely to attrit. Households with an input subsidy card were also significantly less likely to attrit in both samples.

These patterns of characteristics of attrited households are mostly unsurprising. Households with fewer assets and less ties to their land are more likely to attrit, possibly due to migration. Those that are more vulnerable, such as female headed households, are also more likely to attrit. The modest levels of attrition in the BIHS are unlikely to lead to a source of bias in the impact estimates in part because the patterns of attrition are mostly similar in the ZOI and non-ZOI samples. Also, the matching methods used here match treated households in the remaining sample to non-treated households on observables, which removes bias in the estimates by creating an effective statistical comparison group. The greatest threat to the impact estimates from attrition is likely to come in the form of a threat to external validity (representativeness of the underlying population) in that the remaining sample is somewhat different from the original sample. Threats to internal validity (an invalid comparison group) are unlikely to be driven by attrition in this matching analysis.

4. The Impact Evaluation Design

4.1 Overview of identification strategy

We now describe how we will use matching techniques to measure the impact of the Feed the Future initiative in Bangladesh. As with all causal impact evaluations, the fundamental challenge to measuring the impact of Feed the Future is to compare outcomes for households who participated in Feed the Future activities or were exposed to the benefits of Feed the Future investments to the *counterfactual* – what those outcomes would have been if Feed the Future had not undertaken these activities and investments. Of course, households living in the Feed the Future ZOI cannot be observed in the other state of the world, in which Feed the Future activities did not take place, but we can attempt to construct the counterfactual outcomes using comparable data on households living outside the Feed the Future ZOI. We will conduct the matching analysis using the nearest neighbor matching (NNM) form of the covariate matching methods developed by Abadie and Imbens (2006, 2008) in which households that participated in or were exposed to a program are matched directly to nonpartipants based on their observable characteristics. This approach is preferred in this case because we cannot directly observe Feed the Future participation in the BIHS data

set, we can only determine whether the household resides within or outside the Feed the Future ZOI. Next, we describe the approach to estimating the impact of Feed the Future.

4.2 The Matching approach

Matching techniques construct the impact estimates as the average difference in an outcome for each treated household exposed to a program from a weighted average of the outcome in a sample of comparison households in which the weights are based on the degree of similarity between the treated household and the comparison households using observable characteristics. For this analysis, we do not need to observe participation in Feed the Future activities in the BIHS data; we only need to know which households reside in the Feed the Future ZOI, and not whether they participated in Feed the Future activities. In this sense, the households in the Feed the Future ZOI were all "exposed" to the Feed the Future investments. This suggests an "intent-to-treat" interpretation because households in the Feed the Future ZOI had the potential to be treated by Feed the Future either through direct participation, through spillover effects from living near Feed the Future participant households, or through exposure to a Feed the Future system-level market intervention. Although Feed the Future investments in the ZOI in Bangladesh over this period were substantial, the number of households that participated in Feed the Future activities is relatively small as a share of the ZOI population, so for impacts to be measurable it is likely that spillover effects and market- or system-level effects must be large.

We use the covariate matching approach of Abadie and Imbens (2006, 2008) to match households in the Feed the Future ZOI to households not in the ZOI directly on observable characteristics. We note though that it also would be feasible to estimate the impact of Feed the Future using propensity score matching (PSM) methods, including inverse propensity score weighted (IPSW) regression. However, PSM methods tend to perform poorly when a single variable, such as district location, highly predicts program participation. In such settings, there tends to be little overlap in the predicted propensity score between participants and non-participants.

Developing unbiased estimates of impact using the NNM approach of Abadie and Imbens requires two assumptions:

$$E(Y_t^0|X_i, FTF = 1) = E(Y_t^0|X_i, FTF = 0) \qquad \text{(A.1)}$$

and

$$0 < P(X_i) < 1 \qquad \text{(A.2)}$$

Assumption A.1 refers to "conditional mean independence," that conditional on the observable control variables, X_i, households outside the ZOI would have the same outcomes on average as households living in the ZOI if Feed the Future had not been implemented in the ZOI. Assumption A.2 is an "overlap" assumption, that all observations in the comparison group have a strictly positive probability of being in the ZOI. This is consistent with the idea that all areas of Bangladesh were considered during the initial period of targeting, or setting up the location of the Feed the Future activities.

To estimate the NNM model, we recognize that formally the average treatment effect (ATE) is given by

$$ATE = E(Y_t^1 - Y_t^0)$$

and the average treatment effect on the treated (ATET) is

$$ATET = E(Y_t^1 - Y_t^0 | FTF = 1),$$

where Y_t^1 is the expected outcome at time *t* if the treatment is received and Y_t^0 is the expected outcome at time *t* if the treatment is not received, and Feed the Future=1 if the household resides in the Feed the Future ZOI. In practice, each unit or household can only experience one of these potential outcomes by either receiving the treatment or not receiving the treatment. Thus, we construct the counterfactual outcome for households that receive the treatment as the weighted average outcome for those that do not receive it, where the weights, w_i, are based on the degree of similarity in observable characteristics, X: $w_i = w_i(X_i)$.

The NNM method for estimating the potential outcome for the *ith* observation as a function of the observed outcome Y_i is

$$\widehat{Y_{ti}} = \begin{cases} Y_i & if\ t_i = 1 \\ \dfrac{\sum_{j \in NN} w_j Y_j}{\sum_{j \in NN} w_j} & otherwise \end{cases}$$

for $t \in (0,1)$, where NN is the set of nearest neighbors for this observation.

Then the estimated impacts are

$$\widehat{ATE} = \frac{\sum_{i=1}^{n} w_i(\widehat{Y_{1i}} - \widehat{Y_{0i}})}{\sum_{i=1}^{n} w_i} = \frac{\sum_{i=1}^{n} w_i(2t_i - 1)\{1 + K_i\}Y_i}{\sum_{i=1}^{n} w_i} \qquad (3a)$$

and

$$\widehat{ATET} = \frac{\sum_{i=1}^{n} t_i w_i(\widehat{Y_{1i}} - \widehat{Y_{0i}})}{\sum_{i=1}^{n} t_i w_i} = \frac{\sum_{i=1}^{n} \{t_i - (1 - t_i)K_i\}Y_i}{\sum_{i=1}^{n} t_i w_i} \qquad (3b)$$

where K_i is the share of nearest neighbors in the comparison group. Equations (3a) and (3b) are estimated without regression (non-parametrically).

The NNM estimator used here shares many of the useful properties of the PSM estimator: (i) it relies on the same identifying assumptions (A.1 and A.2), (ii) it matches ZOI households to one or more non-ZOI households using baseline characteristics, and (iii) it estimates the average impact as the average of the difference in the outcome for each household in the Feed the Future ZOI from a weighted average of outcomes for matched households outside the ZOI. Unlike PSM, it is not necessary to estimate a model of the predicted probability of receiving the treatment to estimate the NNM model, though we estimate the predicted propensity score model prior to running the NNM model to conduct diagnostic checks, such as whether, for the chosen X variables, the distribution of the predicted probability of receiving the treatment is overlapping in the treated and comparison groups. NNM then calculates the treatment effect in a single model as the average difference in the outcome for each treated household from an imputed weighted

average of outcomes for comparison households where the weights are based on the degree of similarly in the X variables. In this sense, matching estimators like NNM are not "matching" treated and comparison observations, but rather estimating a difference in the outcome from a weighted average of outcomes for similar comparison observations. Other differences between NNM and PSM derive primarily from the rule used to select comparable non-ZOI households and the weights used to construct the difference in weighted average outcomes. As a form of "covariate matching," NNM matches Feed the Future ZOI households to non-ZOI households based directly on the observable characteristics. Each ZOI household is matched to the group of non-ZOI households with the smallest average difference in pre-program characteristics, where this difference is determined using a multi-dimensional metric across all control variables. This approach does not use program eligibility information to create the weights,but does use information from all control variables. A significant advantage of NNM over PSM is that the NNM estimator has analytical standard errors for use in hypothesis tests about significance. PSM, on the other hand, requires estimating bootstrapped standard errors because the regression model must include a predicted propensity score from a first-stage regression. This makes NNM more efficient than PSM, leading to fewer type II errors.[12]

This analysis faces several challenges and limitations. One challenge of estimating the impact of Feed the Future in Bangladesh is that the nonbeneficiary sample must be drawn entirely from households outside the Feed the Future districts, so that geographically-based differences between the Feed the Future ZOI and other parts of Bangladesh may make it challenging to meet the conditional mean independence assumption – that there are no remaining unobserved factors that make households outside the ZOI different from households within the ZOI. To partly address this concern, we present data on coverage of social assistance programs both inside and outside the ZOI, but we do not have information on other programs that households outside the ZOI may have participated in, including government programs or other donor investments. The models also cannot control for differences in agro-ecological conditions or district-level policy differences in and outside the ZOI. The matching models are only able to control for the effect of these differences in context, programs and policies to the extent that they shape the variables used for matching. The extent of the effect of this limitation on the impact estimates cannot be fully known. A challenge for this analysis that differentiates it from most impact evaluations in development economics is that we seek to estimate the impact of Feed the Future on average on all households living in the ZOI, rather than on a clearly identified set of beneficiaries of a specific program or intervention. Because BIHS includes representative samples both inside and outside the ZOI, the sample frame is appropriate to this objective. However, it remains an ambitious objective. Although Feed the Future included major investments, the share of the population directly participating in Feed the Future activities in the ZOI is likely to be small, well below 10 percent (though system- or market-level activities in the ZOI may have reached a higher share of households). Thus, we expect to have less than 200 households, and maybe less than 100 households, out of the 2,040 in the Feed the Future ZOI that participated directly in Feed the Future activities. The estimates we present represent the combined effect of direct participation in Feed the Future interventions along with indirect benefits through spillover effects or system- or market-level interventions.

Before estimating the NNM models, we ran diagnostic tests to assess the quality of the comparison group constructed. Although we do not need to estimate propensity scores for the NNM model, we constructed kernel density graphs of the predicted probability of being in Feed the Future for the households inside and

[12] A type II error occurs in a hypothesis test when one fails to reject the null hypothesis of no effects when that null hypothesis is false (when there actually is an effect).

outside the ZOI to confirm that there is a good overlap of the distributions. This is done to confirm that Assumption A.2 is well satisfied. Second, we tested for differences in average characteristics between households inside and outside the ZOI once the NNM weights are applied. As shown in Tables 2-6, there are many differences in characteristics between households inside and outside the ZOI before the weights are applied, but there should be no statistically significant differences after the weights are applied. This comparison of weighted means helps confirm that assumption A.1 is satisfied. Ultimately, whether these assumptions are well satisfied and the NNM model is robust to alternative specifications or inclusion of X variables is an empirical matter. We include an assessment of the degree to which these assumptions are met and the quality of these models in our analysis at the start of Section 5 below. The process of constructing an effective counterfactual was helped by the fact that the BIHS data contains a rich set of child, maternal/paternal and household control variables. In addition to these, variables that proxy for location effects are included, including those that are measured directly in the community survey and variables constructed from means of the household variables by village (PSU).

To construct the estimates using the BIHS data, we identified a set of X variables for matching from the pre-Feed the Future baseline (2011/12) BIHS data. These control variables must be selected to satisfy two criteria: they must be highly associated with the outcome variable and they must be correlated with the probability of treatment. Satisfying the second criterion is more challenging, but we are helped in this case by the fact that the Feed the Future ZOI was selected in part based on poor nutritional status, food insecurity, and low agricultural productivity.

Appendix Table A.2 presents the initial set of X variables used for the impact analysis. In selecting the X variables for matching, we grouped similar outcomes and use the same set of X variables within outcome groupings, as mentioned above. For example, we used a common set of X variables for Outcome Group 1 related to household wellbeing and empowerment (including household poverty, hunger, food security, and women's empowerment); many of the same household and contextual variables affect these outcomes, although the importance of each X variable will differ by outcome. For child nutrition and infant and young child feeding outcomes (Outcome Group 2), we started with the same set of X variables used for the household wellbeing outcomes and added child and maternal characteristics to explain child HAZ and stunting and child feeding behaviors. Finally, we developed a different set of X variables for Outcome Group 3, the agricultural outcomes (HYV rice variety adoption, rice yields and crop diversity). For this outcome group, we revised the list of household and community variables to include those most relevant to agricultural outcomes and add variables related to land characteristics and farming.

The NNM models were estimated using X variables from the baseline survey and outcome variables from the endline survey. NNM cannot be used with difference-in-differences, except by creating new outcome variables for continuous outcomes that are the difference in outcomes from baseline to endline. This approach is not feasible with binary outcome variables. We prefer to use the same approach for both continuous and binary variables so we only use endline outcome variables.

5. Impact Estimates

5.1 Creating the common support sample

The approach used to develop the specifications for the covariate matching models closely followed the guidance provided in Imbens (2015) and in Rubin (2001). Specifications were developed separately for each of the three outcome groups. Also within outcome group 1, separate specifications were developed for the female and male A-WEAI models. Each outcome group started with the set of X variables listed in Appendix Table A.2. Next, a series of probit regressions were run in which the ZOI dummy variable was regressed on the complete set of X variables to calculate the predicted propensity score. The X variables were then screened and revised using two approaches, one based on the *pscore* routine in STATA and another following the variable selection routine proposed in Imbens (2015). The *pscore* routine was used to identify the common support sample – roughly, the sample in which there is overlap in the distributions of the predicted propensity score (PPS) for the treatment and comparison groups. Estimates of treatment effects based on matching are estimated on a common support sample in order to improve the comparability of the treatment and comparison groups. Setting the common support sample involves identifying observations from the treated sample that have very few observations in the comparison group sample with a similar value of the PPS. Similarly, comparison group observations are flagged for removal from the common support sample if they have no treated observations with a similar predicted propensity score. Typically, the treated observations that fail to satisfy common support are in the upper tail of the PPS distribution; they have a high predicted probability of being in the ZOI and so do not have many close comparison observations according to the PPS. Also, the comparison group observations that fail to satisfy common support are typically in the lower tail of the PPS distribution; they have a low predicted probability of being in the ZOI and so do not have many close treated observations. Observations failing common support were then dropped from the analysis to reduce bias.

The PPS distributions for the ZOI and non-ZOI sample observations that satisfied the common support restriction are presented below in Appendix Figures 2-5. These figures provide support that the specifications satisfy Assumption A.2, the "overlap" assumption. Consider Appendix Figure 2. As expected, the PPS distribution for the ZOI lies to the right of the non-ZOI PPS distribution, as we are predicting similarity with ZOI characteristics. Also, the histograms show that there are no observations at very low or at very high values of the PPS because those observations have been dropped in the common support samples. Each of Appendix Figures 2-5 show substantial overlap in the PPS distributions for the ZOI and non-ZOI samples, suggesting that the overlap assumption is reasonably well satisfied.

5.2 Covariate selection and balance in the matched sample

Another step in the variable selection conducted by *pscore* divides the PPS distribution into intervals of the PPS and estimates t-tests for equality of means for each of the X variables within each of the intervals of the PPS. If any X variables fails the t-test for equality of means in an interval, the interval is split in two and the tests are repeated. This routine is repeated, but if the X variable continues to fail these balancing tests within intervals of the PPS, the variable is flagged as being imbalanced. In these cases, we sought to replace the variable with similar variables and the model was rerun. For example, the variable for household size failed the balancing tests in outcome group 1 for poverty and food security. This variable was then replaced with variables for the number of male and number of female household members and the model was rerun, this time passing the balancing test. In some cases, variables with very little variation in the data failed balancing and were dropped from the model. A dummy variable for Muslim household head is one example because the vast majority of household heads were Muslim. Dropping this variable is also unlikely

to affect estimates for our primary outcomes. This procedure was repeated for each variable that failed balancing until a specification was reached in which very few variables failed balancing by p-score.

Next, we applied the Imbens (2015) routine for variable selection in which groups of variables are included in the probit model in both linear and quadratic terms and with linear interactions with the other variables. Equality of means is then tested for these variables in intervals of the PPS using the normalized difference to identify variables that fail balancing. These variables are removed. The model is then rerun and repeated until the model for that group of variables is stable. This is then repeated for other groups of variables and then all groups of variables are included together and the model is run again to check for balance. We maintained a running tally of variables that failed balancing in either the pscore or Imbens approach. We repeated the specification refinement until ultimately the model no longer failed balancing in the Imbens approach. This model was retained as the final model for that outcome group.

Finally, we estimated the impact of Feed the Future by NNM using the *teffects nnmatch* routine in STATA. After running the NNM model, we conducted a balance check to confirm that the mean of the retained X variables are balanced between the ZOI and non-ZOI samples when the weights from the NNM routine are applied. The results of these balance tests are presented in Appendix Tables B.2-4 below for Outcome Groups 1-3, respectively.[13] In each table, the mean, SD and N are reported for the final set of X variables in the common support sample for the ZOI sample and the non-ZOI sample. The normalized difference is reported for the test of equality of means on the raw data shown in the table and then again using the matched data with weights applied. Most of the normalized differences in the matched data are very small and none of them is larger in absolute value than 0.25, suggesting that these models are balanced and they satisfy Assumption A.1.

5.3 Impact of the Feed the Future initiative

We now present the impact estimates from the covariate matching models. Estimated impacts for Outcome Group 1 are presented in Table 3. Feed the Future had no impact on the poverty rate or the poverty gap. Both estimates are small and neither is statistically significant. Results show that Feed the Future had a beneficial effect in improving household food security, which is demonstrated by a statistically significant reduction in the probability that Feed the Future households experienced any food insecurity (by 8.5 percentage points) or mild food insecurity (by 9.4percentage points). Results also show that Feed the Future increased the probability that a household experienced severe food insecurity (FIES>=7) by 1.8 percentage points on a comparison group mean of 2.2 percent, and this estimate is statistically significant. This estimate is a blemish on otherwise encouraging evidence on the effect of Feed the Future on household food security.

Next, we review the large number of estimates on the impact of Feed the Future on measures of women's and men's empowerment from the A-WEAI indicators.[14] We see positive impacts of Feed the Future on the summary A-WEAI measure of women's empowerment, the 5DE score, of 0.017, along with a substantial statistically significant increase in the probability that a woman is empowered by 5.1 percent (on a comparison group mean of 88.0 percent), shown in Table 3. Feed the Future also increased male's 5DE score (by 0.022), as well as the probability that men are empowered by a statistically significant 5.7

[13] The Outcome Group 1 model is for the poverty and food security outcomes.
[14] A summary of the mean value of these indicators is provided in Appendix Table B.1.

percent (on a comparison group mean of 84.0 percent). Estimated impacts on the gap between a man's and woman's 5DE score and on the probability that the household reached gender parity show small positive effects, but these estimates are not significant at the five percent level. These results are encouraging: the share of women that are empowered is high, higher than for men, and grew as a result of Feed the Future. Empowerment for men grew as well. The WEAI indicators do not present a zero-sum measure of empowerment, so these gains for men do not necessarily come at the cost of a loss in empowerment for women.

We break down the results for women in more details. The results in Table 4 show that Feed the Future reduced the number of agricultural domains in which women have some role in decision-making by 0.2 domains (on a mean of 5.3 domains in the comparison group), a relatively small effect but one that is statistically significant. Feed the Future significantly increased women's participation in groups: the initiative led to a statistically significant 7.5 percentage point increase in the probability that the woman belongs to a group and to an increase in the number of groups in which she is an active member by as statistically significant 0.148. Feed the Future has also increased women's labor supply at the intensive margin, significantly increasing the number of minutes the woman spent on work in the previous day by 15 minutes.

We now turn to the results for men. Feed the Future had a very small but significant effect on the probability that a man solely or jointly owns at least two small or one large asset by 0.8 percentage points. Feed the Future also reduced the number of assets solely or jointly owned by men by a significant 0.169assets. This is a small effect on a comparison group mean of 6.9 assets. Feed the Future increased the number of domains in which men feel that they can make decisions by 0.27 domains and this is statistically significant. This likely reflects an increase in men's economic engagement rather than a loss in control for women as there was no offsetting decline in this indicator for women. Living in the Feed the Future ZOI also increased the group participation for men, with an increase in the probability that a man is part of at least one group by 14.3 percent and an increase in the number of groups in which a man is an active member by 0.24. As with women, Feed the Future increased men's labor supply, significantly increasing the number of minutes men spent on work in the previous day by 22 minutes.

Next, we examine the impact of Feed the Future on indicators of child nutritional status and IYCF practices in Outcome Group 2, shown in Table 5. There is no evidence of impact of Feed the Future on HAZ or stunting prevalence for children age 0-59 months or children age 0-23 months. In addition, Feed the Future had no impact on early initiation of breastfeeding, continued breastfeeding or the probability of a child having minimum dietary diversity. However, Feed the Future did cause a large and significant increase in the probability that a child age 6-23 months consumed iron-rich food by 15.6 percent, compared to a comparison group mean of 38.0 percent of children.

Estimates of the impact of Feed the Future on agriculture, shown in Table 6, are mixed. Feed the Future reduced the probability that the household planted HYV rice by a weakly significant 3.8 percentage points for all households and by a statistically significant 10.3 percentage points for all farming households. The composition of this effect comes about in a small increase in the probability of planting Taus HYV rice (3.5 percentage points), a decrease in the probability of planting Taman HYV rice (7.9 percentage points), and a large decrease in the probability of planting Boro HYV rice (25.4 percentage points). The estimates show that Feed the Future led to a reduction in the share of rice in total

planted area of 13.7 percentage points and a reduction in the share of HYV rice in total planted area of 20.6 percentage points. The former effect may reflect the impact of Feed the Future on crop choice. Indeed, Feed the Future's efforts to promote crop diversification led to a large, significant impact on the number of crops grown, by 0.5 crops per farmer. The reduction in share of cultivated area in HYV rice above the size of the reduction for all rice is disappointing as it suggests a shift toward lower productivity conventional varieties. On rice yields, estimates indicate that Feed the Future reduced rice yields by 0.11 tons per hectare for all rice varieties, but this effect is only weakly significant. However, Feed the Future led to large, significant increases in rice yields for TAus, TAman and Boro HYV varieties and for HYV varieties overall. These results for agricultural production show that Feed the Future played a key role in improving yields of HYV rice varieties and promoting greater crop diversity, during a period in which yields were falling overall.

5.4 Robustness checks

We ran a series of robustness checks to examine the stability of the impact estimates to our model assumptions and empirical specifications.

Trimming the common support sample: In one robustness check, we trimmed the common support samples by removing 5 percent of the remaining observations (2.5% from each tail) and rerunning the models. The common support graphs for these models are shown alongside the common support graphs for the main models in Appendix Figures B.1-B.5. The common support figures with further trimming show a more severe cut of the tails of the PPS distributions suggesting a loss of sample where the overlap assumption had been met, indicating that this is probably not an improvement in the sample. The results of these models are presented in Appendix Tables B.2-B.4. The results are largely quite stable to this change in specification, though the new models show some changes in the empowerment measures and an increase in the probability of continued breastfeeding at one year of 5.5 percent.

Multiple hypothesis correction: Next, we examine whether the main findings based on the impact estimates in Tables 11-13 are affected by adjusting for multiple hypothesis correction. The results presented include a large number of hypothesis tests. Although each hypothesis test is independent, the probability of finding any statistically significant result increases with the number of hypotheses being tested even when no effect truly exists. The reason for this is that, at the 95 percent confidence level (the level used here), there is a 1-in-20 probability of rejecting the null hypothesis by chance and falsely concluding there is a statistically significant effect (also known as type I error in statistics). When many hypotheses are tested, the probability that any of them is significant increases. A commonly used approach to address this concern is to adjust the test statistic for this higher probability of a false rejection of the null hypothesis across all tests. We implemented the correction for multiple hypothesis testing in Anderson (2008) which uses a corrected test statistic, called a sharpened q-value, that accounts for the number of tests conducted. The sharpened q-value is read similarly to a p-value, where $q<0.01$ indicates significance at the 1 percent level, $q<0.05$ indicates significance at the 5 percent level, and $q<0.10$ indicates significance at the 10 percent level.

The impact estimates with sharpened q-values for multiple hypothesis correction are reported in Appendix Tables B.5-B.7. Most of the results are unchanged by this adjustment. The sharpened q-values were constructed by adjusting for the number of hypotheses tested in each outcome group. Although the sharpened q-values are generally higher than the p-values, the result of the hypothesis test is changed for

only a few variables based on the sharpened q-value. In outcome group 1, the decline in the number of assets that are solely or jointly owned by men is only weakly significant with the sharpened q-values (q=0.055), whereas it was significant with the p-values (p=0.032). More importantly, the small increase of 1.3 percentage points in the share of women that are empowered as a result of Feed the Future project is only weakly significant rather than significant after adjusting for multiple hypothesis testing. Similarly, the gap between men's and women's empowerment scores is no longer weakly significant after multiple hypothesis testing correction. In outcome group 2, the effect of Feed the Future in increasing the share of children age 6-23 months who consumed iron-rich foods by 15.4 percentage points is significant at the 5 percent level rather than at the 1 percent level after multiple hypothesis correction, but this does not change our conclusion about that impact estimate. In outcome group 3, none of the conclusions are changed as a result of correcting for multiple hypothesis testing.

Bias due to selection on unobservables: In order to test for the presence of any remaining selection bias even after matching on observables, we implement the Rosenbaum bounds (Rosenbaum 2002). This approach seeks to determine how strongly an unobserved covariate would need to affect selection into treatment status to undermine the main conclusions. The Rosenbaum bounds calculates a statistic, Γ, that is the odds ratio between the value of the outcome in the treatment group and the value of the outcome in the comparison group if the value of the outcome in the treatment group were inflated by an unobserved covariate that was correlated with the outcome, that is, by a selection effect due to unobservables. The Stata command rbounds calculates the p-value for the null hypothesis of zero treatment effect for different levels of the Rosenbaum test statistic, Γ. When Γ=1, there is no unobservables bias in the treatment outcome; if $p<0.05$ when Γ=1, the treatment effect is significant when there is zero bias due to unobservables. As Γ increases, so does the potential contribution of unobservables bias to the difference between the outcome in the treatment and comparison group. The higher the value of Γ at which the p-value is below 0.05, the larger the unobservables bias would need to be to undermine the significant result. That is, the highest value of Γ for which $p<0.05$ tells you the what the odds ratio between the treatment and comparison outcome would need to be for the treatment effect to be driven entirely by bias due to unobservables. Appendix Table B.8 presents the p-values at various levels of the Rosenbaum test statistic for two outcome variables: any food insecurity using FIESand the indicator for whether a child age 6-23 months consumed iron rich foods. Recall from Table 11 that Feed the Future reduced the prevalence of any food insecurity by 8.4 percentage points and this effect was statistically significant at the 1 percent level. The results in Table B.8 show that unobservables bias would have to be large enough to inflate the ratio of the treatment to comparison outcome by more than 40 percent (Γ=1.4) for the measured treatment effect on any food insecurity to no longer be significant. This means that the impact estimate on any food insecurity is relatively unlikely to be undermined by selection on unobservables. For child consumption of iron rich foods, the treatment effect remains significant for Γ=1.4, suggesting that this result is moderately robust to selection on unobservables.

Choice of matching estimator: We also examined the robustness of the results to our choice of matching estimator by estimating all of the matching models again using inverse probability weighted (IPW) regression. IPW regression provides a matching estimator for program treatment effects using a regression model that includes weights equal to the inverse of the PPS in the comparison group and equal to 1 in the treatment group. Appendix tables B.9-B.11 present the impact estimates for outcome groups 1-3, respectively, under the original NNM estimator and under the IPW estimator.

Bias due to differential investment in social assistance: Another potential source of bias could arise if the government or its partners invested more in large-scale social protection programs or other forms of assistance outside the ZOI than inside the ZOI, and if this effect were not controlled for. It is plausible that such investments could be differentially targeted outside the ZOI in response to the fact that Feed the Future is investing heavily in development interventions inside the ZOI. Table 14 presents evidence from the BIHS baseline and endline surveys of the prevalence of participation in a variety of forms of social assistance both inside and outside the ZOI. Overall, differences in receipt of social assistance inside and outside the ZOI are small, but are generally more favorable to ZOI households. The most common form of social assistance in 2018/19 is a stipend for primary school students, which are received by 26.8 percent of households in the ZOI and 25.3 percent of households outside the ZOI. Vulnerable group feeding reaches 13.8 percent of households in the ZOI and only 8.4 percent of households outside the ZOI at endline. The old age allowance is received by 10.6 percent of ZOI households and 8.1 percent of non-ZOI households. This pattern of slightly higher coverage of ZOI households with social assistance is also true for school feeding programs, secondary school stipends, allowances for widowed or destitute women, and several other forms of assistance. This pattern rules out a potential bias in the impact estimates from higher investment in social assistance outsize the ZOI.

Overall, the impact estimates are robust to this change in the matching estimator, as the size of the impact estimates do not change much and the significance level for most outcomes does not change between the NNM and IPW estimator. In outcome group 1 (Table B.9), results based on IPW show that Feed the Future increased the proportion of households below the $1.25 (2005 PPP) poverty line by a significant 3.4 percent. The results for all other outcomes in outcome group 1 were unchanged except for two variables that were only significant at the 10 percent level. In outcome group 2 (Table B.10), there is no difference in significance level for any variable that is significant at the 5 percent level or better. In outcome group 3 (Table B.11), the probability of growing HYV rice for any household declines by a significant 3.6 percent as a result of Feed the Future. Results for other agricultural outcome are largely unchanged in the IPW models, except for rice yields, which decline somewhat more under the IPW model than under NNM, and this result is significant at the 1 percent level.

5.5 Heterogeneity of impact by baseline consumption level

The impact estimates presented thus far are the average impact of Feed the Future on households in the Feed the Future ZOI. However, the impact of the program may differ depending on baseline household characteristics. In particular, an important question is whether the agricultural productivity enhancing interventions in Feed the Future are more successful for less poor households than for the poorest, who may face other constraints in land, credit, labor or skill, for example, that limit their ability to benefit from these interventions. Similarly, the impact of Feed the Future on nutrition and IYCF outcomes may be different for poorer households whose less well-nourished children have a greater ability to respond to nutrition interventions.

In order to test this, we estimated a series of impact models for Outcome Groups 2 and 3 on child nutrition and agriculture in which we separately estimated the impact of Feed the Future on the poorest half of households in the ZOI sample, splitting the sample at the median level of baseline per capita household consumption expenditure. For the purpose of this exposition, we will refer to households below the median per capita household consumption as poor and those above the median per capita household consumption as nonpoor.

We estimated these models by NNM separately in the poor and nonpoor samples. Estimates of the impact of Feed the Future on child nutrition and agriculture by relative poverty status are presented in Tables 15-16. The impact estimates in Table 15 indicate that, among poor households, Feed the Future increased mean HAZ in children under age 5 by 0.179 standard deviations and reduced the prevalence of stunting in children under age 5 by 7.9 percentage points, but each of these estimates is only significant at the 10 percent level. Impact estimates for non-poor households indicate that Feed the Future led to a weakly significant increase in stunting prevalence for children under 5 of 7.8 percent and a weakly significant decrease in stunting prevalence for children under 2 of 10.7 percent. Feed the Future increased the share of children age 6-23 months consuming iron-rich foods in both the poor and nonpoor samples, by a significant 23.8 and 32.3 percent, respectively. Based on these results, we cannot conclusively state that Feed the Future improved nutritional status of children under 5 in poor households because the estimates are not significant at conventional levels ($p<0.05$). Still, these results are suggestive, as they do not indicate a precisely estimated null effect.

The heterogeneous results for agriculture outcomes in Table 16 show that the pattern of negative average treatment effects on the adoption of HYV rice varieties occur both among the poor and nonpoor household in the sample. Also, the share of rice and HYV rice in planted area falls as a result of Feed the Future in both the poor and nonpoor samples. The negative effects on overall rice yield are somewhat larger in the poor sample and are only weakly significant, though this is probably driven by the reduction in sample size. The reduction in yield in the nonpoor sample is much smaller and is not significant. However, the significant impacts of Feed the Future on rice yields for TAman and Boro in the overall sample area maintained in the poor sample and, for Boro rice, in the nonpoor sample, indicating that the beneficial effects are shared by poor and nonpoor households alike.

To check the sensitivity of these heterogenous impact estimates to the choice of matching estimator, we also estimated these models by IPW. Those results are reported in Appendix Tables B.12 and B.13. We find that the results are not entirely robust to the change in matching estimator. Using IPW, stunting prevalence is reduced by Feed the Future by a significant 10.3 percentage points in the poor sample and the prevalence of consumption of iron-rich foods by children is not affected. In the nonpoor sample, there is a significant increase in the probability of continued breastfeeding at one year of 3.3 percentage points and the probability of a child consuming iron-rich foods increases by a significant 13.0 percentage points.

6. Conclusion

This paper presents evidence on the impact of the USAID-funded Feed the Future initiative in Bangladesh. This represents the first time that a country-level Feed the Future program has undergone a quantitative impact evaluation to measure its overall effects. The impact estimates presented should be interpreted as the average impact on households living in the Feed the Future ZOI, regardless of their direct participation in specific Feed the Future activities. Thus, the effects measured are general equilibrium effects, including spillover effects and market- and systems-level effects within the ZOI from all Feed the Future activities.

The results of this evaluation show that Feed the Future had no impact on the poverty rate in the ZOI, which was the program's highest overall objective. The analysis cannot clearly identify the reasons that the poverty objective was not met, but we can speculate about possible reasons. One plausible contributing factor is that the Feed the Future initiative, while large, was not large enough. Of the 28

million people living in the ZOI, it is likely that Feed the Future directly reached less than 10 percent of them. Thus, the initiative would have needed to have very large indirect effects across the ZOI, to have a measurable impact on poverty. Another possible explanation is that the program was insufficiently targeted to poor households to be able to have a measurable impact on poverty alleviation. Many beneficiaries of the program were not living below the poverty line, so the benefits that they obtained from their participation would not contribute to the poverty reduction objective unless their participation had positive spillover effects on the poor. A third reason is that many Feed the Future investments are designed to improve outcomes over time, such as through increased productivity in agriculture, but may not have as significant an effect on poverty as social assistance programs that are more narrowly focused on the objective of reducing poverty in the short run. It is possible that, had all Feed the Future investments been targeted directly and effectively to the poor as cash transfers, there may have been a measurable effect of the program on poverty, but that beneficial effect may have been short lived. This explains why graduation model programs have become popular alternatives to cash transfer programs for achieving longer-run impacts on poverty, though the model's cost is a challenge. Overall, we expect that Feed the Future may have had larger effects on poverty if it had been targeted more directly to poor households and if the program also included regular consumption support, in the form of cash or food transfers, in addition to its many investments in other sectors. Recent evidence suggests that blended models of targeted cash transfers with complementary investments in livelihoods, health and gender equality may hold promise as a modality that could be effective at achieving meaningful poverty reduction.

Despite the lack of impact on poverty, Feed the Future led to a substantial improvement in food security, reducing the mean FIES by 0.15 and reducing the probability of any food insecurity by 8.4 percentage points. The small increase in severe food insecurity caused by Feed the Future is not easily explained; it may reflect some weakness in the response to an acute shock in the Feed the Future ZOI. Measures of women's and men's empowerment show that Feed the Future increased group participation by both women and men. Also, women and men increased their labor supply at the intensive margin, increasing minutes worked. Other dimensions of male empowerment also increased. Feed the Future had a small, positive impact on the summary A-WEAI measure of women's empowerment across 5 domains, increasing the share of women that are rated as empowered by this indicator by 4.6 percent (on a comparison group mean of 88.0 percent). This empowerment effect is fairly robust: the significance of this estimate survived a correction for multiple hypothesis testing. The probability that men are empowered across these 5 domains increased by a statistically significant 5.7 percent (on a comparison group mean of 84.0 percent) and this result was also robust to multiple hypothesis testing.

There were no significant impacts of Feed the Future on child nutritional status, including child stunting. Although we cannot definitively explain this null result, many of the most effective direct interventions to reduce child stunting combine a substantial income boost for poor households with intensive nutrition behavior change communication trainings and improved linkages to health services and WASH (Ruel and Alderman 2013). The pathways operating in Feed the Future, through improved household food security, greater agricultural productivity, more crop diversification and strengthened women's empowerment may all contribute to an enabling environment for stunting reduction, but the size of these beneficial effects would need to be much larger, particularly on income benefits accruing to households with children at risk of malnutrition.

Within IYCF outcomes, Feed the Future increased the probability that a child age 6-23 months consumed iron-rich foods by 15.4 percentage points. Estimated impacts of Feed the Future on agriculture suggest that the program reduced adoption of HYV rice varieties and further contributed to the trend in declining yields in rice farming observed in the national BIHS sample. Estimates showed no significant heterogeneity in impacts on child nutrition outcomes or agriculture outcomes across the poorest half and least poor half of the sample.

This impact evaluation shows that, despite a lack of impacts on poverty and child stunting, Feed the Future had effects on household food security, agricultural productivity, crop diversification, women's empowerment and iron source foods in child diets. Moreover, these measured effects are found in the sectors in which Feed the Future focused its investments, showing that donor priorities are able to shape outcomes in recipient countries. At the same time, the limited size of the effects suggests that, to be more effective, future programs will need to be better targeted, and should combine direct transfers with the most promising complementary programs identified here.

References

Abadie, A., and G. Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* 74(1): 235-267.

Abadie, A., and G. Imbens. 2008. "On the Failure of the Bootstrap for Matching Estimators." *Econometrica* 76(6): 1537-1557.

Ahmed, A.A., M.M. Bakhtiar, D.O. Gilligan, and G. Thai. 2020. Research Protocol: Feed the Future Zone of Influence Analysis. Washington DC: International Food Policy Research Institute.

Anderson, Michael L. 2008. Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association* 103(484):1481-1495.

Arndt, C., Jones, S., & Tarp, F. 2015. Assessing Foreign Aid's Long-Run Contribution to Growth and Development. *World Development* 69, 6-18.

Balagtas, J. V., Bhandari, H., & Cabrera, E. 2014. Did the Comprehensive Agrarian Reform Program Improve Agricultural Productivity in the Philippines?. *Food Policy*, 44, 200-209.

Barrett, C. B., Carter, M. R., & Chavas, J. P. 2012. The Economics of Poverty Traps and Persistent Poverty: An Asset-Based Approach. *Journal of Development Studies*, 49(7), 976-990.

Bazzi, S., & Clemens, M. A. 2013. Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth. *American Economic Journal: Macroeconomics*, 5(2), 152-186.

Berthélemy, J. C. 2006. Bilateral Donors' Interest vs. Recipients' Development Motives in Aid Allocation: Do All Donors Behave the Same?. *Review of Development Economics*, 10(2), 179-194.

Bhutta, Z. A., Das, J. K., Rizvi, A., Gaffey, M. F., Walker, N., Horton, S., Webb, P., Lartey, A., & Black, R. E. 2013. Evidence-based interventions for improvement of maternal and child nutrition: what can be done and at what cost?. *The Lancet*, 382(9890), 452-477.

Burnside, C., & Dollar, D. 2000. Aid, Policies, and Growth. *American Economic Review*, 90(4), 847-868.

Chandy, L., & Kharas, H. 2011. Why Can't We Just Give Money to the Poor?. Brookings Institution.

Clemens, M. A., Radelet, S., Bhavnani, R. R., & Bazzi, S. 2012. Counting Chickens When They Hatch: Timing and the Effects of Aid on Growth. *Economic Journal*, 122(561), 590-617.

Collier, P. 2007. The Bottom Billion: Why the Poorest Countries are Failing and What Can Be Done About It. Oxford University Press.

De Janvry, A., & Sadoulet, E. 2010. Agricultural Growth and Poverty Reduction: Additional Evidence. *The World Bank Research Observer*, 25(1), 1-20.

Dietrich, S. 2013. Bypass or Engage? Explaining Donor Delivery Tactics in Aid Allocation. *International Studies Quarterly*, 57(4), 698-712.

Dorosh, P., & Thurlow, J. 2014. Can Cities or Towns Drive African Development? Economy-wide Analysis for Ethiopia and Uganda. *World Development*, 63, 113-123.

Dreher, A., Fuchs, A., Parks, B., Strange, A. M., & Tierney, M. J. 2017. Aid, China, and Growth: Evidence from a New Global Development Finance Dataset. AidData Working Paper #46. Williamsburg, VA: AidData.

Easterly, W. 2006. The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good. Oxford University Press.

Feed the Future. 2020. Feed the Future Bangladesh Zone of Influence Endline Assessment April 2013–March 2020.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. 2016. Impact Evaluation in Practice. World Bank Publications.

Headey, D., Hoddinott, J., & Park, S. 2012. Drivers of Nutritional Change in Four South Asian Countries: A Dynamic Panel Analysis. IFPRI Discussion Paper 01195.

Heckman, J.J., H. Ichimura, and P.E. Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program." *Review of Economic Studies* 64:605-654.

_____. 1998. "Matching as an Econometric Evaluation Estimator." *Review of Economic Studies* 65:261-294.

Heckman, J., H. Ichimura, J.A. Smith, and P.E. Todd. 1998. "Characterizing Selection Bias Using Experimental Data." *Econometrica* 66:1017-1098.

Hirano, K., Imbens, G., and Ridder, G. 2003. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica* 71 (4): 1161-1189.

Hoddinott, J., Headey, D., & Dereje, M. 2013. Cows, Missing Milk Markets, and Nutrition in Rural Ethiopia. *Journal of Development Studies*, 49(8), 1010-1023.

IFPRI. 2020. Impact Evaluation: Bangladesh AVC. Endline report on impacts associated with the Bangladesh agricultural value chains project. Dhaka, Bangladesh: International Food Policy Research Institute.

Imbens, G. W. 2015. Matching Methods in Practice: Three Examples. *Journal of Human Resources* 50(2):373-419.

Imbens, G. W., and Rubin, D. B. 2015. Causal inference: For statistics, social, and biomedical sciences an introduction, in *Causal Inference: For Statistics, Social, and Biomedical Sciences: an Introduction*.

Khandker, S. R., Koolwal, G. B., & Samad, H. A. 2012. Handbook on Impact Evaluation: Quantitative Methods and Practices. World Bank Publications.

Khandker, S. R., Samad, H. A., & Ali, R. 2010. Does Access to Finance Matter in Microenterprise Growth? Evidence from Bangladesh. *World Development*, 38(12), 1799-1812.

Leroy, J. L., Ruel, M., Frongillo, E. A., Harris, J., & Ballard, T. J. 2016. Measuring the Food Access Dimension of Food Security: A Critical Review and Mapping of Indicators. *Food and Nutrition Bulletin*, 37(2), 167-184.

Malapit, H., C. Pinkstaff, K. Sproule, C. Kovarik, A. Quisumbing and R. Meinzen-Dick. 2017. The Abbreviated Women's Empowerment in Agriculture Index (A-WEAI). IFPRI Discussion Paper 01647. Washington DC: International Food Policy Research Institute.

Minten, B., Tamru, S., Kuma, T., & Nyarko, Y. 2013. Structure and Performance of Ethiopia's Coffee Export Sector. IFPRI Discussion Paper 01254.

Moyo, D. 2009. Dead Aid: Why Aid Is Not Working and How There Is a Better Way for Africa. Farrar, Straus and Giroux.

National Institute of Population Research and Training (NIPORT), Mitra and Associates, and ICF International. 2016. Bangladesh Demographic and Health Survey 2014. Dhaka, Bangladesh, and Rockville, Maryland, USA: NIPORT, Mitra and Associates, and ICF International.

OECD. (2019). Development aid drops in 2018, especially to neediest countries. OECD. https://www.oecd.org/dac/financing-sustainable-development/development-finance-data/ODA-2018-detailed-summary.pdf

Peterman, A., Behrman, J. A., & Quisumbing, A. R. (2014). A Review of Empirical Evidence on Gender Differences in Nonland Agricultural Inputs, Technology, and Services in Developing Countries. IFPRI Discussion Paper 013.

Rosenbaum, P. 2002. *Observational Studies*, Second edition. Springer Series in Statistics, Springer-Verlag, New York.

Rubin, D.B. 2001. Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation. *Health Services & Outcomes Research Methodology* 2:169–188.

Ruel, Marie, Harold Alderman and the Maternal and Child Nutrition Study Group. 2013. Nutrition-sensitive interventions and programmes: how can they help to accelerate progress in improving maternal and child nutrition? *Lancet* 382: 536-51.

USAID. (2022, November 10). Bangladesh. [Website]. https://www.feedthefuture.gov/country/bangladesh/

USAID Feed the Future Annual Report 2023: https://www.usaid.gov/news-information/press-releases/jun-22-2023-united-states-announces-260-million-additional-funding-address-global-food-security-crisis

Table 1: Summary statistics on covariates for the matching analysis, for ZOI and non-ZOI samples, 2011/12

	ZOI Sample			Non-ZOI Sample			P-value	Normalized difference
	Mean	SD	N	Mean	SD	N		
Household head's age	45.828	13.888	2,064	44.642	13.999	4,380	0.001	0.069
Household head's age squared	2,292.988	1,362.348	2,064	2,188.800	1,350.714	4,380	0.004	0.059
Household head is female	0.148	0.355	2,064	0.170	0.376	4,380	0.028	0.014
Household head's years of schooling	3.637	4.008	2,064	3.081	3.836	4,380	0.000	0.166
Household head's years of schooling squared	29.287	44.036	2,064	24.205	39.978	4,380	0.000	0.139
Household head's age * years of schooling	160.184	194.104	2,064	128.041	175.351	4,380	0.000	0.185
Spouse's years of schooling	3.370	3.494	1,722	2.866	3.438	3,518	0.000	0.152
Dependency ratio (# dependents/# working people)	0.821	0.713	2,031	0.876	0.740	4,327	0.006	-0.088
Total number of male members	2.096	1.126	2,064	2.142	1.254	4,380	0.159	-0.067
Total number of female members	2.194	1.108	2,064	2.320	1.229	4,380	0.000	-0.099
Household head's occupation is agriculture	0.466	0.499	2,064	0.417	0.493	4,380	0.000	0.061
Household head's occupation is agricultural wage employment	0.132	0.338	2,064	0.131	0.338	4,380	0.955	-0.013
Household head's occupation is non-ag (wage emp or business)	0.266	0.442	2,064	0.281	0.449	4,380	0.199	-0.027
Total land owned (hectare)	0.269	0.504	2,064	0.213	0.480	4,380	0.000	0.137
Total tropical livestock units owned	0.384	0.389	2,064	0.338	0.379	4,380	0.000	0.105

Household wealth index (PCA)	0.197	1.953	2,064	0.087	1.983	4,380	0.037	0.081
HH productive asset index (PCA)	0.154	1.766	2,064	0.103	1.641	4,380	0.253	0.001
House has more than 1 room	0.290	0.454	2,064	0.191	0.393	4,380	0.000	0.208
Improved outer wall (concrete or tin)	0.644	0.479	2,064	0.612	0.487	4,380	0.014	0.089
Improved floor (concrete or tin)	0.107	0.309	2,064	0.118	0.322	4,380	0.187	-0.001
Improved latrine (flush or w/o flush, water sealed)	0.299	0.458	2,064	0.306	0.461	4,380	0.581	0.008
Improved drinking water (piped or tubewell)	0.857	0.350	2,064	0.873	0.333	4,380	0.077	-0.015
HH had at least 1 negative shock in past 5 years	0.599	0.490	2,064	0.509	0.500	4,380	0.000	0.189
HH had at least 1 positive shock in past 5 years	0.181	0.385	2,064	0.274	0.446	4,380	0.000	-0.249
Village level average HH wealth index (PCA)	0.203	0.822	2,064	0.087	0.912	4,380	0.000	0.158
Village level proportion of HHs with improved latrine	0.298	0.183	2,064	0.306	0.188	4,380	0.100	0.026
Village level proportion of HHs with improved water	0.857	0.217	2,064	0.873	0.198	4,380	0.003	-0.032
Village has at least 1 primary school	0.803	0.398	2,064	0.713	0.452	4,380	0.000	0.230
Village has at least 1 clinic	0.312	0.463	2,064	0.226	0.418	4,380	0.000	0.163
Village's average distance to nearest town < 10km	0.620	0.485	2,064	0.623	0.485	4,380	0.823	-0.026
Improved roof (concrete or tin)	0.914	0.280	2,064	0.966	0.182	4,380	0.000	-0.226
Village level average years of schooling of HHHs	3.636	1.423	2,064	3.081	1.344	4,380	0.000	0.431
Village has at least 1 NGO program	0.936	0.246	2,064	0.937	0.242	4,380	0.772	-0.045
Village has improved roads	0.389	0.488	2,064	0.374	0.484	4,380	0.252	0.033
HH has ag. input subsidy card	0.131	0.337	2,064	0.072	0.259	4,380	0.000	0.202

Presence of irrigation	0.600	0.490	2,064	0.518	0.500	4,380	0.000	0.159
Planted more than 1 crops	0.424	0.494	2,064	0.333	0.471	4,380	0.000	0.183
Number of plots owned and operated	1.500	2.515	2,064	1.215	2.440	4,380	0.000	0.118
Total cultivated land (hectare)	0.428	0.651	2,064	0.364	0.652	4,380	0.000	0.096
Village has at least 1 ag supplier of fertilizer	0.547	0.498	2,064	0.509	0.500	4,380	0.005	0.024
Village has at least 1 ag supplier of seeds	0.429	0.495	2,064	0.469	0.499	4,380	0.003	-0.118
Village has at least 1 ag supplier of pesticide	0.523	0.500	2,064	0.497	0.500	4,380	0.046	0.014
Village has canal irrigation	0.212	0.409	2,064	0.166	0.372	4,343	0.000	0.157
Mother's age	26.722	5.856	762	26.965	5.758	1,832	0.331	-0.037
Mother's years of schooling	6.961	3.485	762	6.275	3.449	1,832	0.000	0.186
Mother's BMI	22.921	4.012	743	22.203	3.829	1,776	0.000	0.182
Mother's dietary diversity score	4.891	1.012	728	4.924	1.095	1,760	0.485	-0.031
Child's age in month	29.511	17.965	780	29.869	17.073	1,887	0.627	-0.019
Child's sex	0.524	0.500	780	0.517	0.500	1,887	0.719	0.017
Child's birth order	2.078	1.184	780	2.316	1.379	1,887	0.000	-0.174
Number of older siblings	0.878	0.926	780	1.051	1.102	1,887	0.000	-0.159
Presence of older sister	0.379	0.486	780	0.399	0.490	1,887	0.360	-0.031
Child had diarrhea in the past 4 weeks	0.071	0.256	780	0.062	0.241	1,887	0.416	0.034

Notes: The household wealth index and household productive asset index are each constructed using Principal Components Analysis (PCA). Data are from the 2011/12 round of the Bangladesh Integrated Household Survey. SD refers to standard deviations. P-values are from a t-test for of equality of means of the outcome across the ZOI and non-ZOI samples. The normalized difference for each outcome is defined as the ratio of the ZOI sample mean minus the non-ZOI sample mean to the square root of one half of the sum of the ZOI sample and non-ZOI sample variances.

Table 2: Tests for selective attrition in household and village characteristics, in baseline ZOI and non-ZOI samples

	ZOI Sample					Non-ZOI Sample				
	Attrited HHs		Non-attrited HHs			Attrited HHs		Non-attrited HHs		
	Mean/SD	N	Mean/SD	N	p-value	Mean/SD	N	Mean/SD	N	p-value
Household head's age	43.724	127	45.828	2,064	0.103	41.278	291	44.642	4,380	0.000
	(17.525)		(13.888)			(15.689)		(13.999)		
Household head's age squared	2,216.543	127	2,292.988	2,064	0.547	1,949.216	291	2,188.800	4,380	0.004
	(1,767.048)		(1,362.348)			(1,554.660)		(1,350.714)		
Household head is female	0.378	127	0.148	2,064	0.000	0.254	291	0.170	4,380	0.000
	(0.487)		(0.355)			(0.436)		(0.376)		
Household head's years of schooling	3.772	127	3.637	2,064	0.714	3.045	291	3.081	4,380	0.875
	(4.162)		(4.008)			(3.791)		(3.836)		
Household head's years of schooling squared	31.409	127	29.287	2,064	0.601	23.595	291	24.205	4,380	0.800
	(49.401)		(44.036)			(38.024)		(39.978)		
Household head's age * years of schooling	145.929	127	160.184	2,064	0.420	115.354	291	128.041	4,380	0.230
	(180.925)		(194.104)			(163.680)		(175.351)		
Spouse's years of schooling	4.152	79	3.370	1,722	0.053	2.709	213	2.866	3,518	0.516

	(3.749)		(3.494)			(3.416)		(3.438)		
Dependency ratio (# dependents/# working people)	0.946	115	0.821	2,031	0.070	0.983	275	0.876	4,327	0.020
	(0.796)		(0.713)			(0.772)		(0.740)		
Total number of male members	1.417	127	2.096	2,064	0.000	1.625	291	2.142	4,380	0.000
	(0.996)		(1.126)			(0.990)		(1.254)		
Total number of female members	1.913	127	2.194	2,064	0.005	1.976	291	2.320	4,380	0.000
	(1.008)		(1.108)			(0.973)		(1.229)		
Household head's occupation is agriculture	0.394	127	0.466	2,064	0.112	0.289	291	0.417	4,380	0.000
	(0.491)		(0.499)			(0.454)		(0.493)		
Household head's occupation is agricultural wage employment	0.071	127	0.132	2,064	0.046	0.189	291	0.131	4,380	0.005
	(0.258)		(0.338)			(0.392)		(0.338)		
Household head's occupation is non-ag (wage emp or business)	0.260	127	0.266	2,064	0.889	0.302	291	0.281	4,380	0.428
	(0.440)		(0.442)			(0.460)		(0.449)		
Total land owned (hectare)	0.145	127	0.269	2,064	0.006	0.123	291	0.213	4,380	0.002
	(0.317)		(0.504)			(0.291)		(0.480)		
Total tropical livestock units owned	0.231	127	0.384	2,064	0.000	0.159	291	0.338	4,380	0.000
	(0.356)		(0.389)			(0.287)		(0.379)		

Household wealth index (PCA)	-0.760	127	0.197	2,064	0.000	-0.813	291	0.087	4,380	0.000
	(1.938)		(1.953)			(2.043)		(1.983)		
HH productive asset index (PCA)	-0.515	127	0.154	2,064	0.000	-0.621	291	0.103	4,380	0.000
	(1.470)		(1.766)			(1.051)		(1.641)		
House has more than 1 room	0.134	127	0.290	2,064	0.000	0.082	291	0.191	4,380	0.000
	(0.342)		(0.454)			(0.276)		(0.393)		
Improved outer wall (concrete or tin)	0.543	127	0.644	2,064	0.022	0.598	291	0.612	4,380	0.631
	(0.500)		(0.479)			(0.491)		(0.487)		
Improved floor (concrete or tin)	0.055	127	0.107	2,064	0.065	0.107	291	0.118	4,380	0.562
	(0.229)		(0.309)			(0.309)		(0.322)		
Improved latrine (flush or w/o flush, water sealed)	0.197	127	0.299	2,064	0.014	0.306	291	0.306	4,380	0.996
	(0.399)		(0.458)			(0.462)		(0.461)		
Improved drinking water (piped or tubewell)	0.811	127	0.857	2,064	0.153	0.787	291	0.873	4,380	0.000
	(0.393)		(0.350)			(0.410)		(0.333)		
HH had at least 1 negative shock in past 5 years	0.606	127	0.599	2,064	0.876	0.598	291	0.509	4,380	0.003
	(0.491)		(0.490)			(0.491)		(0.500)		
HH had at least 1 positive shock in past 5 years	0.173	127	0.181	2,064	0.831	0.210	291	0.274	4,380	0.017

	(0.380)		(0.385)			(0.408)		(0.446)		
Village level average HH wealth index (PCA)	-0.025	127	0.161	2,064	0.014	-0.193	291	0.046	4,380	0.000
	(0.881)		(0.824)			(0.957)		(0.915)		
Village level proportion of HHs with improved latrine	0.268	127	0.293	2,064	0.127	0.303	291	0.306	4,380	0.813
	(0.172)		(0.181)			(0.194)		(0.186)		
Village level proportion of HHs with improved water	0.824	127	0.856	2,064	0.112	0.840	291	0.870	4,380	0.017
	(0.227)		(0.218)			(0.215)		(0.201)		
Village has at least 1 primary school	0.748	127	0.803	2,064	0.135	0.739	291	0.713	4,380	0.349
	(0.436)		(0.398)			(0.440)		(0.452)		
Village has at least 1 clinic	0.315	127	0.312	2,064	0.945	0.220	291	0.226	4,380	0.824
	(0.466)		(0.463)			(0.415)		(0.418)		
Village's average distance to nearest town < 10km	0.575	127	0.620	2,064	0.308	0.598	291	0.623	4,380	0.392
	(0.496)		(0.485)			(0.491)		(0.485)		
Improved roof (concrete or tin)	0.874	127	0.914	2,064	0.121	0.945	291	0.966	4,380	0.068
	(0.333)		(0.280)			(0.228)		(0.182)		
Village level average years of schooling of HHHs	3.566	127	3.651	2,064	0.509	2.997	291	3.084	4,380	0.284

	(1.393)		(1.401)			(1.412)		(1.340)		
Village has at least 1 NGO program	0.898	127	0.936	2,064	0.096	0.942	291	0.937	4,380	0.777
	(0.304)		(0.246)			(0.235)		(0.242)		
Village has improved roads	0.354	127	0.389	2,064	0.436	0.344	291	0.374	4,380	0.296
	(0.480)		(0.488)			(0.476)		(0.484)		
HH has ag. input subsidy card	0.055	127	0.131	2,064	0.013	0.034	291	0.072	4,380	0.014
	(0.229)		(0.337)			(0.182)		(0.259)		
Presence of irrigation	0.354	127	0.600	2,064	0.000	0.351	291	0.518	4,380	0.000
	(0.480)		(0.490)			(0.478)		(0.500)		
Planted more than 1 crops	0.220	127	0.424	2,064	0.000	0.165	291	0.333	4,380	0.000
	(0.416)		(0.494)			(0.372)		(0.471)		
Number of plots owned and operated	0.748	127	1.500	2,064	0.001	0.381	291	1.215	4,380	0.000
	(1.447)		(2.515)			(1.005)		(2.440)		
Total cultivated land (hectare)	0.194	127	0.428	2,064	0.000	0.133	291	0.364	4,380	0.000
	(0.406)		(0.651)			(0.299)		(0.652)		
Village has at least 1 ag supplier of fertilizer	0.433	127	0.547	2,064	0.013	0.460	291	0.509	4,380	0.111
	(0.497)		(0.498)			(0.499)		(0.500)		
Village has at least 1 ag supplier of seeds	0.307	127	0.429	2,064	0.007	0.405	291	0.469	4,380	0.036

	(0.463)		(0.495)			(0.492)		(0.499)		
Village has at least 1 ag supplier of pesticide	0.433	127	0.523	2,064	0.048	0.440	291	0.497	4,380	0.061
	(0.497)		(0.500)			(0.497)		(0.500)		
Village has canal irrigation	0.205	127	0.212	2,064	0.851	0.118	288	0.166	4,343	0.033
	(0.405)		(0.409)			(0.323)		(0.372)		

Notes: The household wealth index and household productive asset index are each constructed using Principal Components Analysis (PCA). Data are from the 2011/12 round of the Bangladesh Integrated Household Survey. SD refers to standard deviations. P-values are from a t-test for of equality of means of the outcome across the ZOI and non-ZOI samples.

Table 3: Impact estimates for aggregate indices

	Treatment effect (SE) from NNM	P-value	Sharpened q-value	N	Comparison mean
Poverty and food security					
Standardized weighted poverty index	0.044	0.186	0.148	5,159	0.000
	(0.033)				
HH Food Insecurity Experience Scale (raw score: 0-8)	-0.144**	0.028	0.032	5,159	1.552
	(0.066)				
Women's empowerment					

Standardized weighted women's A-WEAI index - all indicators	0.076**	0.022	0.027	4,660	0.000
	(0.033)				
Woman's five domains of empowerment (5DE) score	0.017***	0.006	0.011	4,697	0.875
	(0.006)				
Woman is empowered	0.051***	0.001	0.004	4,697	0.637
	(0.016)				
Standardized weighted men's A-WEAI index - all indicators	0.120***	0.001	0.003	4,090	-0.000
	(0.036)				
Man's five domains of empowerment (5DE) score	0.022***	0.002	0.006	4,117	0.833
	(0.007)				
Man is empowered	0.057***	0.002	0.005	4,117	0.558
	(0.018)				
Gap between man's and woman's 5DE scores within household	0.006	0.169	0.137	3,953	0.030
	(0.004)				
Household reached gender parity	0.027*	0.053	0.051	4,697	0.750
	(0.014)				
Agricultural production					
Standardized weighted agricultural production index	-0.125***	0.006	0.003	3,014	-0.000

	(0.045)

Notes: Estimates from the common support sample are selected by the matching model. The standardized weighted poverty index is constructed from (i) a dummy variable for being a poor household based on daily per capita consumption less than $1.25 PPP and (ii) poverty gap where poverty line at $1.25 PPP. The standardized weighted WEAI indices for women and men are constructed using both binary and continuous A-WEAI indicators. The standardized weighted agricultural production index is aggregated from rice production indicators and number of crop grown for all farmers. All standardized weighted index outcomes are aggregated following the procedure in Anderson (2008) such that the comparison group has mean of 0 and standard deviation of 1. Sharpened q-values in column 4 follow Benjamini et al. (2006) to control the false discovery rate. Stars denote statistical significance as follows: * $p<0.10$; ** $p<0.05$; *** $p<0.01$

Table 4: Impact estimates for poverty and women's empowerment outcomes

	Treatment effect (SE) from NNM	P-value	Sharpened q-value	N	Comparison mean
Poverty and food security					
Poor HH: daily per cap consumption expenditure < $1.25 PPP	0.023*	0.091	0.077	5,159	0.207
	(0.013)				
Poverty gap where poverty line = $1.25 PPP	0.275	0.410	0.246	5,159	3.900
	(0.333)				
Any Food Insecurity (FIES score >=1)	-0.085***	0.000	0.001	5,159	0.511
	(0.016)				
Mild Food Insecurity (1<=FIES<=3)	-0.094***	0.000	0.001	5,159	0.354
	(0.015)				
Moderate Food Insecurity (4<=FIES<=6)	-0.010	0.399	0.244	5,159	0.135

	(0.011)				
Severe Food Insecurity (FIES>=7)	0.018***	0.005	0.011	5,159	0.022
	(0.007)				
Women's A-WEAI indicators					
Has some input in decisions/feels like can make decisions in at least 1 domain	-0.004	0.422	0.224	4,660	0.984
	(0.004)				
No of ag domains resp has some input in decision making	-0.190**	0.027	0.032	4,660	5.339
	(0.086)				
Solely/jointly owns at least 2 small or 1 large asset, incl land	0.008	0.217	0.155	4,660	0.950
	(0.006)				
No of assets that are solely/jointly owned by resp	0.051	0.362	0.244	4,660	3.928
	(0.056)				
Has some input in income dec or feels like can make decisions	-0.005	0.469	0.286	4,660	0.955
	(0.007)				
No of domains resp has some input in income decisions/feels can make decision	0.043	0.431	0.247	4,660	3.926
	(0.055)				
Jointly makes at least 1 decision regarding at least 1 source of credit	0.020	0.217	0.155	4,660	0.613
	(0.016)				

No. of accessible sources where solely/jointly participated in credit decisions	0.033	0.289	0.196	4,660	0.906
	(0.031)				
Resp is part of at least 1 group	0.075***	0.000	0.001	4,660	0.589
	(0.016)				
No of groups where resp is an active member	0.148***	0.000	0.001	4,660	0.742
	(0.026)				
Resp didn't work more than 10.5 hours per day	-0.022	0.159	0.158	4,660	0.747
	(0.015)				
No of minutes spent on work	15.523***	0.008	0.014	4,660	517.629
	(5.846)				
Men's A-WEAI indicators					
Has some input in decisions/feels like can make decisions in at least 1 domain	-0.001	0.931	0.472	4,090	0.962
	(0.007)				
No of ag domains resp has some input in decision making	0.043	0.636	0.344	4,090	5.526
	(0.092)				
Solely/jointly owns at least 2 small or 1 large asset, incl land	-0.008**	0.011	0.016	4,090	0.998
	(0.003)				
No of assets that are solely/jointly owned by resp	-0.169**	0.019	0.023	4,090	6.864

	(0.072)				
Has some input in income dec or feels like can make decisions	0.005	0.339	0.244	4,090	0.971
	(0.005)				
No of domains resp has some input in income decisions/feels can make decision	0.269***	0.000	0.001	4,090	4.832
	(0.065)				
Jointly makes at least 1 decision regarding at least 1 source of credit	0.003	0.833	0.444	4,090	0.762
	(0.015)				
No. of accessible sources where solely/jointly participated in credit decisions	0.033	0.308	0.224	4,090	1.140
	(0.032)				
Resp is part of at least 1 group	0.143***	0.000	0.001	4,090	0.261
	(0.018)				
No of groups where resp is an active member	0.239***	0.000	0.001	4,090	0.315
	(0.026)				
Resp didn't work more than 10.5 hours per day	-0.019	0.283	0.215	4,090	0.661
	(0.018)				
No of minutes spent on work	22.238***	0.003	0.009	4,090	519.118
	(7.582)				

Notes: Estimates from the common support sample are selected by the matching model. Sharpened q-values in column 4 follow Benjamini et al. (2006) to control the false discovery rate. Stars denote statistical significance as follows: * $p<0.10$; ** $p<0.05$; *** $p<0.01$

Table 5: Impact estimates for child outcomes

	Treatment effect (SE) from NNM	P-value	Sharpened q-value	N	Comparison mean
Under 5 children: Length/height-for-age Z-score	0.070	0.344	0.224	1,930	-1.489
	(0.074)				
Under 5 children: Stunting (HAZ < -2)	-0.019	0.506	0.301	1,930	0.338
	(0.029)				
Under 2 children: Length/height-for-age Z-score	0.138	0.300	0.210	756	-1.332
	(0.133)				
Under 2 children: Stunting (HAZ < -2)	-0.066	0.153	0.126	756	0.304
	(0.046)				
IYCF practice: early initiation of breastfeeding for child 0-23mons	-0.013	0.772	0.431	762	0.778
	(0.044)				
IYCF practice: continued breastfeeding at 1 year for child 12-23mons	0.031	0.233	0.171	338	0.939
	(0.026)				
IYCF practice: Minimum dietary diversity (>=5 out of 8 food groups) for child 6-	-0.053	0.306	0.210	576	0.282
	(0.052)				
IYCF practice: Consumption of iron-rich food for child 6-23mons	0.156***	0.007	0.012	576	0.391
	(0.057)				

Notes: Estimates from the common support sample are selected by the matching model. Sharpened q-values in column 4 follow Benjamini et al. (2006) to control the false

discovery rate. Stars denote statistical significance as follows: * $p<0.10$; ** $p<0.05$; *** $p<0.01$

Table 6: Impact estimates for agricultural production outcomes

	Treatment effect (SE) from NNM	P-value	Sharpened q-value	N	Comparison mean
Planted HYV rice - all HHs	-0.038**	0.014	0.019	5,122	0.516
	(0.016)				
Planted HYV rice - all farmers	-0.103***	0.000	0.001	3,014	0.905
	(0.015)				
Planted HYV rice: TAus variety - all farmers	0.035***	0.006	0.011	3,014	0.054
	(0.012)				
Planted HYV rice: TAman variety - all farmers	-0.079***	0.000	0.001	3,014	0.646
	(0.021)				
Planted HYV rice: Boro variety - all farmers	-0.254***	0.000	0.001	3,014	0.744
	(0.020)				
Percentage of all rice out of total crop area - all farmers	-0.137***	0.000	0.001	3,014	0.818
	(0.014)				
Percentage of HYV rice out of total crop area - all farmers	-0.206***	0.000	0.001	3,014	0.786
	(0.015)				
Rice production/ha in metric ton - rice farmers	-0.111**	0.025	0.030	2,767	3.163

	(0.049)				
HYV TAus rice production/ha in metric ton - HYV TAus rice farmers	-0.067	0.794	0.438	149	2.437
	(0.256)				
HYV TAman rice production/ha in metric ton - HYV TAman rice farmers	0.155***	0.004	0.009	1,868	2.731
	(0.054)				
HYV Boro rice production/ha in metric ton - HYV Boro rice farmers	0.411***	0.000	0.001	1,955	3.687
	(0.059)				
Rice variety production/ha in metric ton excluding local varieties	0.110**	0.030	0.032	2,635	3.247
	(0.051)				
Number of crops grown - all farmers	0.459***	0.000	0.001	3,014	1.663
	(0.058)				

Notes: Estimates from the common support sample are selected by the matching model. Sharpened q-values in column 4 follow Benjamini et al. (2006) to control the false discovery rate. Stars denote statistical significance as follows: * $p<0.10$; ** $p<0.05$; *** $p<0.01$

Appendix A: Outcome variables and matching variables for the impact analysis

Table A.1: Outcome variables for the Feed the Future Impact Analysis

Prevalence of poverty • Share of sample below $1.25 per capita expenditure per day (2005 PPP)
Food insecurity • Food Insecurity Experience Scale (FIES) index, which categorizes households into three levels: any food insecurity, moderate to severe food insecurity, severe food insecurity
Women's empowerment: A-WEAI 5 domains and 6 indicators in both binary (adequacy) and continuous (number of sub-domains) measures: • production • resource (with 2 indicators) • income • leadership • time
Prevalence of stunting among children under 5 • HAZ score • Stunting (HAZ < -2)
Child feeding behaviors • Early initiation of breastfeeding • Continued breastfeeding at 1 year • Introduction of solid, semi-solid, or soft food • Minimum dietary diversity • Consumption of iron-rich food
Application of improved technologies • Adoption of HYV rice

Rice yield • Rice yield in last 1 year (metric ton per hectare), including by rice type
Crop diversity • Number of crops grown in the last year

Table A.2: Matching variables for the Feed the Future Impact Analysis

Outcome Group 1: Household expenditure, hunger, food security, women's empowerment
Household demographics and sector of employment • household head age; household head is female; household head is married; household head years of schooling; household head is muslim; • years of schooling of spouse of household head • household size; household dependency ratio; household male to female ratio; number of disabled household members • household head occupation is agriculture; household head occupation is agricultural wage employment; household head occupation is non-ag wage employment; household head occupation is non-ag business • total land owned (hectare); number of large livestock owned; number of small/medium livestock owned; household wealth index • household owned dwelling; number of rooms in the house; household has improved wall; household has improved roof; household has improved floor; household has improved latrine; household has improved water • number of negative shocks experienced in the past 5 years; number of positive shocks experienced in the past 5 years (at baseline)
Community characteristics • average wealth PCA, average years in school for household heads, proportion of households with improved latrine, proportion of household with improved water

- presence of at least one primary school, presence of at least one health clinic, presence of at least one NGO development program, most common road type is paved road
- average distance to the nearest town (km)

Outcome Group 2: child HAZ, child stunting, IYCF practices

All matching variables from Outcome Group 1, plus…

Child characteristics

- age, sex, birth order, birth spacing (age difference from nearest older sibling)
- children under 2: children's dietary diversity
- children under 2: ever breastfed
- number of older siblings
- presence of older sister
- any diarrhea in the past 4 weeks

Maternal characteristics

- mother's age
- mother's years of schooling
- mother's BMI
- mother's dietary diversity score

Outcome Group 3: high yielding variety rice adoption, rice yields

Household demographics and sector of employment

- household head age; household head is female; household head is married; household head years of schooling; household head is muslim;
- years of schooling of spouse of household head
- household size; household dependency ratio; household male to female ratio; number of disabled household members
- household head occupation is agriculture; household head occupation is agricultural wage employment; household head occupation is non-ag wage employment; household head occupation is non-ag business
- number of large livestock owned; number of small/medium livestock owned; household wealth index; household productive asset index

- household owned dwelling; number of rooms in the house; household has improved wall; household has improved roof; household has improved floor; household has improved latrine; household has improved water
- number of negative shocks experienced in the past 5 years; number of positive shocks experienced in the past 5 years (at baseline)
- household had agricultural extension agent visit in the last 12 months
- household had agricultural input subsidy card

Land and farming characteristics

- presence of irrigation; number of plots owned; number of crops grown
- total land cultivated (hectare)
- household used any fertilizer or pesticides on plots
- household used draft animal or machinery on plots
- average flood depth of cultivated land (ft)
- the dominant soil type of cultivated land is clay-loam

Community characteristics

- average wealth PCA, average years in school for household heads, proportion of households with improved latrine, proportion of household with improved water
- presence of at least one primary school, proportion of households indicating access to extension agents, presence of at least one NGO development program, most common road type is paved road
- average distance to the nearest town (km)
- presence of at least one supplier of fertilizer, presence of at least one supplier of seeds, presence of at least one supplier of pesticide
- average agricultural labor wage

Appendix B: Additional results

Appendix Table B.1: Trend in outcomes between the ZOI and non-ZOI samples, 2011/12-2018/19

	Baseline 2011-2012			Endline 2018-2019			P-value	Normalized difference
	Mean	SD	N	Mean	SD	N		
Poor HH: daily per cap consumption expenditure < $1.25 PPP	0.374	0.484	2,064	0.215	0.411	2,064	0.000	-0.354
HH answered yes to one of hunger questions	0.136	0.343	2,064	0.064	0.246	2,064	0.000	-0.240
Planted HYV rice - all farmers	0.800	0.400	1,238	0.807	0.395	1,185	0.662	0.018
Planted HYV rice: TAus variety - all farmers	0.120	0.326	1,238	0.083	0.276	1,185	0.002	-0.125
Planted HYV rice: TAman variety - all farmers	0.393	0.489	1,238	0.562	0.496	1,185	0.000	0.344
Planted HYV rice: Boro variety - all farmers	0.556	0.497	1,238	0.492	0.500	1,185	0.002	-0.128
Percentage of all rice out of total crop area - all farmers	0.665	0.361	1,238	0.680	0.349	1,185	0.285	0.043
Rice production/ha in metric ton - rice farmers	2.827	1.218	1,091	3.081	1.166	1,066	0.000	0.213
HYV TAus rice production/ha in metric ton - rice farmers	2.155	0.936	149	2.809	0.874	98	0.000	0.723
HYV TAman rice production/ha in metric ton - rice farmers	2.526	0.759	486	2.844	0.917	666	0.000	0.378
HYV Boro rice production/ha in metric ton - rice farmers	3.780	1.167	688	4.149	0.966	583	0.000	0.345
Rice production/ha in mton, excluding local varieties	3.136	1.199	986	3.367	1.098	956	0.000	0.201
Number of crops grown - all farmers	2.294	1.619	1,238	2.096	1.420	1,185	0.001	-0.130
Under 5 children: Length/height-for-age Z-score	-1.782	1.543	777	-1.348	1.277	758	0.000	0.306
Under 5 children: Stunting (HAZ < -2)	0.430	0.495	777	0.301	0.459	758	0.000	-0.270

Under 2 children: Length/height-for-age Z-score	-1.520	1.657	313	-1.143	1.410	305	0.002	0.245
Under 2 children: Stunting (HAZ < -2)	0.335	0.473	313	0.246	0.431	305	0.014	-0.198
Non-ZOI sample								
Poor HH: daily per cap consumption expenditure < $1.25 PPP	0.351	0.477	4,380	0.193	0.395	4,380	0.000	-0.361
HH answered yes to one of hunger questions	0.081	0.273	4,380	0.067	0.250	4,380	0.011	-0.054
Planted HYV rice - all farmers	0.883	0.322	2,271	0.894	0.308	2,264	0.234	0.035
Planted HYV rice: TAus variety - all farmers	0.097	0.296	2,271	0.056	0.229	2,264	0.000	-0.156
Planted HYV rice: TAman variety - all farmers	0.528	0.499	2,271	0.623	0.485	2,264	0.000	0.193
Planted HYV rice: Boro variety - all farmers	0.754	0.431	2,271	0.726	0.446	2,264	0.028	-0.065
Percentage of all rice out of total crop area - all farmers	0.807	0.302	2,271	0.809	0.310	2,264	0.877	0.005
Rice production/ha in metric ton - rice farmers	2.833	1.112	2,118	3.147	0.987	2,066	0.000	0.299
HYV TAus rice production/ha in metric ton - rice farmers	1.871	0.759	220	2.351	0.902	126	0.000	0.575
HYV TAman rice production/ha in metric ton - rice farmers	2.148	1.162	1,200	2.702	0.863	1,411	0.000	0.541
HYV Boro rice production/ha in metric ton - rice farmers	3.611	1.078	1,713	3.685	1.027	1,643	0.040	0.071
Rice production/ha in mton, excluding local varieties	3.025	1.131	2,013	3.230	0.970	2,024	0.000	0.195
Number of crops grown - all farmers	1.850	1.393	2,271	1.643	1.149	2,264	0.000	-0.162
Under 5 children: Length/height-for-age Z-score	-1.849	1.564	1,946	-1.488	1.286	1,821	0.000	0.252
Under 5 children: Stunting (HAZ < -2)	0.479	0.500	1,946	0.332	0.471	1,821	0.000	-0.302
Under 2 children: Length/height-for-age Z-score	-1.554	1.632	809	-1.317	1.368	721	0.002	0.157

Under 2 children: Stunting (HAZ < -2)	0.394	0.489	809	0.288	0.453	721	0.000	-0.224

Notes: Data are from the 2011/12 and 2018/19 rounds of the Bangladesh Integrated Household Survey. For the variables related to rice farming, the number of observations (N) is the number of households cultivating rice in that survey round and subsample. SD refers to standard deviations. P-values are from a t-test of equality of means of the outcome across the ZOI and non-ZOI samples. The normalized difference for each outcome is defined as the ratio of the endline mean minus the baseline mean to the square root of one half of the sum of the baseline and endline variances.

Appendix Table B.2: Balance table for covariates in outcome group 1, by ZOI and non-ZOI samples

	ZOI sample			Non-ZOI sample			Normalized difference (raw)	Normalized difference (matched)
	Mean	SD	N	Mean	SD	N		
Household head's age	46.543	13.305	1,704	45.590	13.562	3,455	0.071	-0.013
Household head's age squared	2,343.214	1,318.244	1,704	2,262.319	1,323.736	3,455	0.061	-0.026
Household head is female	0.005	0.073	1,704	0.005	0.068	3,455	0.009	-0.000
Household head's years of schooling	3.662	4.060	1,704	3.041	3.867	3,455	0.157	0.022
Household head's years of schooling squared	29.887	45.325	1,704	24.203	41.006	3,455	0.132	0.004
Household head's age * years of schooling	167.925	202.523	1,704	133.280	184.359	3,455	0.179	0.024
Spouse's years of schooling	3.403	3.496	1,704	2.909	3.448	3,455	0.142	0.045
Dependency ratio (# dependents/# working people)	0.753	0.588	1,704	0.802	0.590	3,455	-0.084	-0.075
Total number of male members	2.303	1.040	1,704	2.370	1.196	3,455	-0.060	-0.040
Total number of female members	2.231	1.113	1,704	2.342	1.252	3,455	-0.094	-0.092
Household head's occupation is agriculture	0.477	0.500	1,704	0.443	0.497	3,455	0.066	0.017

Household head's occupation is agricultural wage employment	0.148	0.355	1,704	0.153	0.360	3,455	-0.013	0.004
Household head's occupation is non-ag (wage emp or business)	0.302	0.459	1,704	0.317	0.465	3,455	-0.032	-0.007
Total land owned (hectare)	0.292	0.498	1,704	0.222	0.457	3,455	0.147	0.034
Total tropical livestock units owned	0.425	0.393	1,704	0.383	0.386	3,455	0.110	0.022
Household wealth index (PCA)	0.333	1.907	1,704	0.181	1.952	3,455	0.079	-0.040
HH productive asset index (PCA)	0.320	1.833	1,704	0.308	1.679	3,455	0.007	-0.082
House has more than 1 room	0.305	0.461	1,704	0.216	0.412	3,455	0.204	0.019
Improved outer wall (concrete or tin)	0.638	0.481	1,704	0.599	0.490	3,455	0.081	0.041
Improved floor (concrete or tin)	0.103	0.304	1,704	0.106	0.307	3,455	-0.010	-0.043
Improved latrine (flush or w/o flush, water sealed)	0.297	0.457	1,704	0.292	0.455	3,455	0.011	-0.069
Improved drinking water (piped or tubewell)	0.870	0.336	1,704	0.876	0.330	3,455	-0.017	0.050
HH had at least 1 negative shock in past 5 years	0.613	0.487	1,704	0.518	0.500	3,455	0.192	0.122
HH had at least 1 positive shock in past 5 years	0.177	0.382	1,704	0.275	0.447	3,455	-0.236	-0.166
Village level average HH wealth index (PCA)	0.211	0.806	1,704	0.087	0.899	3,455	0.145	0.040
Village level proportion of HHs with improved latrine	0.300	0.185	1,704	0.295	0.187	3,455	0.025	-0.006
Village level proportion of HHs with improved water	0.867	0.208	1,704	0.874	0.198	3,455	-0.035	0.029
Village has at least 1 primary school	0.817	0.386	1,704	0.724	0.447	3,455	0.225	0.168
Village has at least 1 clinic	0.312	0.464	1,704	0.239	0.426	3,455	0.165	0.040

Village's average distance to nearest town < 10km	0.615	0.487	1,704	0.627	0.484	3,455	-0.025	-0.016

Notes: Estimates from the comparison and beneficiary groups selected by the matching model.

Appendix Table B.3: Balance table for covariates in outcome group 2, by ZOI and non-ZOI samples

	ZOI sample			Non-ZOI sample			Normalized difference (raw)	Normalized difference (matched)
	Mean	SD	N	Mean	SD	N		
Household head's age	46.100	14.832	617	44.392	14.956	1,328	0.121	0.038
Household head's age squared	2,344.872	1,430.405	617	2,194.188	1,434.989	1,328	0.112	0.028
Household head is female	0.006	0.080	617	0.005	0.067	1,328	0.026	-0.009
Household head's years of schooling	3.684	4.069	617	2.919	3.663	1,328	0.188	0.059
Household head's years of schooling squared	30.102	46.289	617	21.927	36.528	1,328	0.190	0.036
Household head's age * years of schooling	164.404	197.726	617	124.553	177.216	1,328	0.206	0.051
Spouse's years of schooling	3.464	3.606	617	3.016	3.406	1,328	0.118	0.025
Dependency ratio (# dependents/# working people)	0.664	0.567	617	0.697	0.553	1,328	-0.055	-0.094
Total number of male members	2.355	1.162	617	2.443	1.363	1,328	-0.066	-0.058
Total number of female members	2.371	1.191	617	2.462	1.373	1,328	-0.069	-0.080
Household head's occupation is agriculture	0.506	0.500	617	0.448	0.497	1,328	0.113	0.099
Household head's occupation is agricultural wage employment	0.123	0.329	617	0.146	0.353	1,328	-0.062	-0.026

Household head's occupation is non-ag (wage emp or business)	0.295	0.456	617	0.303	0.460	1,328	-0.013	-0.029
Total land owned (hectare)	0.344	0.548	617	0.255	0.612	1,328	0.150	0.054
Total tropical livestock units owned	0.399	0.385	617	0.378	0.387	1,328	0.052	0.056
Household wealth index (PCA)	0.371	1.926	617	0.245	2.029	1,328	0.063	0.001
HH productive asset index (PCA)	0.397	1.960	617	0.307	1.762	1,328	0.052	-0.036
House has more than 1 room	0.363	0.481	617	0.226	0.418	1,328	0.311	0.180
Improved outer wall (concrete or tin)	0.692	0.462	617	0.596	0.491	1,328	0.203	0.086
Improved roof (concrete or tin)	0.927	0.260	617	0.956	0.204	1,328	-0.125	-0.004
Improved floor (concrete or tin)	0.099	0.299	617	0.104	0.305	1,328	-0.018	-0.095
Improved latrine (flush or w/o flush, water sealed)	0.280	0.450	617	0.309	0.462	1,328	-0.058	-0.077
Improved drinking water (piped or tubewell)	0.857	0.350	617	0.855	0.352	1,328	0.010	0.105
HH had at least 1 negative shock in past 5 years	0.587	0.493	617	0.517	0.500	1,328	0.147	0.121
HH had at least 1 positive shock in past 5 years	0.152	0.360	617	0.231	0.422	1,328	-0.195	-0.208
Village level average HH wealth index (PCA)	0.301	0.929	617	0.195	1.137	1,328	0.100	0.141
Village level average years of schooling of HHHs	3.532	1.765	617	3.031	1.737	1,328	0.274	0.187
Village level proportion of HHs with improved latrine	0.279	0.205	617	0.313	0.246	1,328	-0.146	-0.107
Village level proportion of HHs with improved water	0.856	0.245	617	0.858	0.226	1,328	-0.004	0.129
Village has at least 1 primary school	0.801	0.400	617	0.721	0.449	1,328	0.192	0.174
Village has at least 1 clinic	0.310	0.463	617	0.238	0.426	1,328	0.161	0.020

Village has at least 1 NGO program	0.925	0.263	617	0.934	0.249	1,328	-0.032	0.070
Village has improved roads	0.376	0.485	617	0.354	0.478	1,328	0.048	0.105
Village's average distance to nearest town < 10km	0.600	0.490	617	0.615	0.487	1,328	-0.029	0.018
Mother's age	27.057	5.883	617	27.271	5.640	1,328	-0.040	-0.072
Mother's years of schooling	6.848	3.549	617	6.119	3.555	1,328	0.201	0.120
Mother's BMI	22.991	4.123	617	22.129	3.726	1,328	0.223	0.144
Mother's dietary diversity score	4.867	1.025	617	4.865	1.110	1,328	0.003	0.033
Child's age in month	28.853	17.712	617	29.402	16.995	1,328	-0.033	-0.022
Child's sex	0.517	0.500	617	0.525	0.500	1,328	-0.010	-0.002
Child's birth order	2.164	1.208	617	2.383	1.337	1,328	-0.169	-0.157
Number of older siblings	0.925	0.923	617	1.137	1.090	1,328	-0.207	-0.172
Presence of older sister	0.405	0.491	617	0.428	0.495	1,328	-0.041	-0.046
Child had diarrhea in the past 4 weeks	0.078	0.268	617	0.062	0.242	1,328	0.065	-0.013

Notes: Estimates from the comparison and beneficiary groups selected by the matching model.

Appendix Table B.4: Balance table for covariates in outcome group 3, by ZOI and non-ZOI samples

	ZOI sample			Non-ZOI sample			Normalized difference (raw)	Normalized difference (matched)
	Mean	SD	N	Mean	SD	N		
Household head's age	46.543	13.305	1,704	45.591	13.555	3,418	0.071	-0.004

Household head's age squared	2,343.214	1,318.244	1,704	2,262.256	1,321.867	3,418	0.061	-0.017
Household head is female	0.005	0.073	1,704	0.004	0.066	3,418	0.013	0.000
Household head's years of schooling	3.662	4.060	1,704	3.022	3.861	3,418	0.162	0.032
Household head's years of schooling squared	29.887	45.325	1,704	24.031	41.019	3,418	0.135	0.013
Household head's age * years of schooling	167.925	202.523	1,704	132.575	184.839	3,418	0.182	0.030
Spouse's years of schooling	3.403	3.496	1,704	2.898	3.442	3,418	0.146	0.069
Dependency ratio (# dependents/# working people)	0.753	0.588	1,704	0.800	0.587	3,418	-0.081	-0.064
Total number of male members	2.303	1.040	1,704	2.368	1.184	3,418	-0.059	-0.033
Total number of female members	2.231	1.113	1,704	2.331	1.222	3,418	-0.086	-0.073
Household head's occupation is agriculture	0.477	0.500	1,704	0.446	0.497	3,418	0.062	-0.013
Household head's occupation is agricultural wage employment	0.148	0.355	1,704	0.152	0.359	3,418	-0.013	-0.003
Household head's occupation is non-ag (wage emp or business)	0.302	0.459	1,704	0.316	0.465	3,418	-0.029	0.037
Total land owned (hectare)	0.292	0.498	1,704	0.223	0.470	3,418	0.141	0.025
Total tropical livestock units owned	0.425	0.393	1,704	0.384	0.386	3,418	0.107	-0.001
Household wealth index (PCA)	0.333	1.907	1,704	0.173	1.951	3,418	0.083	-0.048
HH productive asset index (PCA)	0.320	1.833	1,704	0.307	1.677	3,418	0.007	-0.110
House has more than 1 room	0.305	0.461	1,704	0.216	0.412	3,418	0.204	0.059
Improved outer wall (concrete or tin)	0.638	0.481	1,704	0.595	0.491	3,418	0.089	0.010
Improved floor (concrete or tin)	0.103	0.304	1,704	0.102	0.303	3,418	0.002	-0.046

Improved latrine (flush or w/o flush, water sealed)	0.297	0.457	1,704	0.292	0.455	3,418	0.010	-0.060
Improved drinking water (piped or tubewell)	0.870	0.336	1,704	0.874	0.332	3,418	-0.012	0.043
HH had at least 1 negative shock in past 5 years	0.613	0.487	1,704	0.523	0.500	3,418	0.182	0.110
HH had at least 1 positive shock in past 5 years	0.177	0.382	1,704	0.276	0.447	3,418	-0.238	-0.188
HH has ag. input subsidy card	0.151	0.358	1,704	0.087	0.281	3,418	0.199	0.029
Presence of irrigation	0.665	0.472	1,704	0.590	0.492	3,418	0.158	-0.007
Planted more than 1 crops	0.478	0.500	1,704	0.390	0.488	3,418	0.179	0.040
Number of plots owned and operated	1.699	2.675	1,704	1.391	2.633	3,418	0.116	0.017
Total cultivated land (hectare)	0.491	0.682	1,704	0.426	0.689	3,418	0.095	-0.032
Village level average HH wealth index (PCA)	0.211	0.806	1,704	0.083	0.902	3,418	0.150	0.068
Village level proportion of HHs with improved latrine	0.300	0.185	1,704	0.295	0.187	3,418	0.027	-0.006
Village level proportion of HHs with improved water	0.867	0.208	1,704	0.873	0.199	3,418	-0.028	0.014
Village has at least 1 primary school	0.817	0.386	1,704	0.726	0.446	3,418	0.218	0.136
Village has at least 1 clinic	0.312	0.464	1,704	0.242	0.428	3,418	0.158	0.030
Village has improved roads	0.408	0.492	1,704	0.395	0.489	3,418	0.027	-0.022
Village's average distance to nearest town < 10km	0.615	0.487	1,704	0.627	0.484	3,418	-0.025	-0.020
Village has at least 1 ag supplier of fertilizer	0.553	0.497	1,704	0.543	0.498	3,418	0.020	-0.055
Village has at least 1 ag supplier of seeds	0.437	0.496	1,704	0.498	0.500	3,418	-0.121	-0.068
Village has at least 1 ag supplier of pesticide	0.533	0.499	1,704	0.528	0.499	3,418	0.010	-0.053
Village has canal irrigation	0.215	0.411	1,704	0.156	0.363	3,418	0.153	0.086

Notes: Estimates from the comparison and beneficiary groups selected by the matching model.

Appendix Table B.5: Trend in participation in social safety net programs between the ZOI and non-ZOI samples, 2011/12-2018/19

	Baseline 2011-2012			Endline 2018-2019			P-value	Normalized difference
	Mean	SD	N	Mean	SD	N		
Received assistance from stipend for primary students	0.246	0.431	2,064	0.268	0.443	2,064	0.109	0.050
Received assistance from vulnerable group feeding	0.041	0.198	2,064	0.138	0.345	2,064	0.000	0.346
Received assistance from old age allowance	0.058	0.234	2,064	0.106	0.307	2,064	0.000	0.174
Received assistance from school feeding program	0.043	0.203	2,064	0.062	0.242	2,064	0.005	0.087
Received assistance from stipend for secondary/higher secondary/female students	0.066	0.248	2,064	0.050	0.219	2,064	0.033	-0.066
Received assistance from allowances for widowed, deserted and destitute women	0.017	0.131	2,064	0.031	0.175	2,064	0.003	0.091
Received assistance from vulnerable group development	0.027	0.163	2,064	0.026	0.158	2,064	0.771	-0.009
Received assistance from allowances for the financially insolvent disabled	0.004	0.066	2,064	0.015	0.120	2,064	0.001	0.105
Received assistance from pension program for retired govt employees and families	0.000	0.000	2,064	0.010	0.100	2,064	0.000	0.143
Non-ZOI sample								
Received assistance from stipend for primary students	0.222	0.416	4,380	0.254	0.435	4,380	0.001	0.074
Received assistance from vulnerable group feeding	0.068	0.252	4,380	0.083	0.276	4,380	0.009	0.056

Received assistance from old age allowance	0.046	0.209	4,380	0.082	0.274	4,380	0.000	0.146
Received assistance from school feeding program	0.011	0.103	4,380	0.041	0.198	4,380	0.000	0.191
Received assistance from stipend for secondary/higher secondary/female students	0.050	0.219	4,380	0.030	0.170	4,380	0.000	-0.106
Received assistance from allowances for widowed, deserted and destitute women	0.012	0.110	4,380	0.026	0.159	4,380	0.000	0.099
Received assistance from vulnerable group development	0.019	0.136	4,380	0.011	0.106	4,380	0.004	-0.062
Received assistance from allowances for the financially insolvent disabled	0.002	0.043	4,380	0.011	0.104	4,380	0.000	0.115
Received assistance from pension program for retired govt employees and families	0.000	0.000	4,380	0.007	0.084	4,380	0.000	0.119

Notes: Data are from the 2011/12 and 2018/19 rounds of the Bangladesh Integrated Household Survey. For the variables related to rice farming, the number of observations (N) is the number of households cultivating rice in that survey round and subsample. SD refers to standard deviations. P-values are from a t-test of equality of means of the outcome across the ZOI and non-ZOI samples. The normalized difference for each outcome is defined as the ratio of the endline mean minus the baseline mean to the square root of one half of the sum of the baseline and endline variances.

Appendix Table B.6: Impact results by poor and non-poor households using NNM - Outcome group 2

	Only poor households			Only non-poor households		
	Treatment effect (SE)	N	Comparison mean	Treatment effect (SE)	N	Comparison mean
Under 5 children: Length/height-for-age Z-score	0.179*	946	-1.682	-0.124	984	-1.315
	(0.097)			(0.121)		
Under 5 children: Stunting (HAZ < -2)	-0.079*	946	0.389	0.078*	984	0.292
	(0.041)			(0.040)		

Under 2 children: Length/height-for-age Z-score	0.183	362	-1.558	0.141	394	-1.153
	(0.156)			(0.193)		
Under 2 children: Stunting (HAZ < -2)	0.014	362	0.345	-0.101	394	0.271
	(0.066)			(0.064)		
IYCF practice: early initiation of breastfeeding for child 0-23mons	0.057	365	0.810	-0.063	397	0.753
	(0.066)			(0.060)		
IYCF practice: continued breastfeeding at 1 year for child 12-23mons	0.047**	169	0.957	0.026	203	0.924
	(0.019)			(0.053)		
IYCF practice: Minimum dietary diversity (>=5 out of 8 food groups) for child 6-	0.048	273	0.235	-0.045	303	0.321
	(0.082)			(0.057)		
IYCF practice: Consumption of iron-rich food for child 6-23mons	0.265***	273	0.374	0.348***	303	0.405
	(0.097)			(0.085)		

Notes: Estimates from the common support sample selected by the matching model. The poverty measure is defined as below and above median daily per capita consumption expenditure at baseline. * $p<0.10$; ** $p<0.05$; *** $p<0.01$

Appendix Table B.7: Impact results by poor and non-poor households using NNM - Outcome group 3

	Only poor households			Only non-poor households		
	Treatment	N	Comparison	Treatment	N	Comparison

	effect (SE)		mean	effect (SE)		mean
Planted HYV rice - all HHs	-0.058***	2,670	0.524	-0.022	2,452	0.506
	(0.022)			(0.024)		
Planted HYV rice - all farmers	-0.109***	1,592	0.899	-0.092***	1,422	0.913
	(0.022)			(0.025)		
Planted HYV rice: TAus variety - all farmers	0.026	1,592	0.058	0.036**	1,422	0.050
	(0.018)			(0.018)		
Planted HYV rice: TAman variety - all farmers	-0.110***	1,592	0.654	-0.070**	1,422	0.637
	(0.029)			(0.031)		
Planted HYV rice: Boro variety - all farmers	-0.264***	1,592	0.723	-0.214***	1,422	0.766
	(0.029)			(0.030)		
Percentage of all rice out of total crop area - all farmers	-0.148***	1,592	0.819	-0.123***	1,422	0.816
	(0.020)			(0.022)		
Percentage of HYV rice out of total crop area - all farmers	-0.218***	1,592	0.782	-0.202***	1,422	0.790
	(0.022)			(0.023)		
Rice production/ha in metric ton - rice farmers	-0.147**	1,469	3.158	-0.073	1,298	3.168
	(0.067)			(0.077)		
HYV TAman rice production/ha in metric ton - HYV TAman rice farmers	0.231***	1,010	2.738	0.145*	858	2.721
	(0.076)			(0.085)		

HYV Boro rice production/ha in metric ton - HYV Boro rice farmers	0.231***	1,003	3.719	0.520***	952	3.654
	(0.089)			(0.087)		
Number of crops grown - all farmers	0.451***	1,592	1.608	0.537***	1,422	1.723
	(0.070)			(0.105)		

Notes: Estimates from the common support sample selected by the matching model. The poverty measure is defined as below and above median daily per capita consumption expenditure at baseline. * $p<0.10$; ** $p<0.05$; *** $p<0.01$

Appendix Table B.8: Impact estimates for poverty and women's empowerment outcomes - Remove coastal districts

	Treatment effect (SE) from NNM	N	Comparison mean
Poverty and food security			
Standardized weighted poverty index	0.015	4,549	0.000
	(0.037)		
Poor HH: daily per cap consumption expenditure < $1.25 PPP	0.021	4,549	0.207
	(0.016)		
Poverty gap where poverty line = $1.25 PPP	-0.149	4,549	3.889
	(0.340)		
HH Food Insecurity Experience Scale (raw score: 0-8)	-0.303***	4,549	1.550
	(0.078)		

Any Food Insecurity (FIES score >=1)	-0.103***	4,549	0.511
	(0.020)		
Mild Food Insecurity (1<=FIES<=3)	-0.077***	4,549	0.355
	(0.018)		
Moderate Food Insecurity (4<=FIES<=6)	-0.034***	4,549	0.133
	(0.013)		
Severe Food Insecurity (FIES>=7)	0.008	4,549	0.023
	(0.007)		
Women's A-WEAI indicators			
Has some input in decisions/feels like can make decisions in at least 1 domain	-0.001	4,326	0.984
	(0.004)		
No of ag domains resp has some input in decision making	-0.198**	4,326	5.341
	(0.091)		
Solely/jointly owns at least 2 small or 1 large asset, incl land	0.003	4,326	0.950
	(0.007)		
No of assets that are solely/jointly owned by resp	0.111*	4,326	3.934
	(0.061)		
Has some input in income dec or feels like can make decisions	-0.004	4,326	0.955

	(0.008)		
No of domains resp has some input in income decisions/feels can make decision	0.017	4,326	3.925
	(0.058)		
Jointly makes at least 1 decision regarding at least 1 source of credit	0.028	4,326	0.613
	(0.018)		
No. of accessible sources where solely/jointly participated in credit decisions	0.058*	4,326	0.902
	(0.035)		
Resp is part of at least 1 group	0.082***	4,326	0.589
	(0.017)		
No of groups where resp is an active member	0.145***	4,326	0.742
	(0.029)		
Resp didn't work more than 10.5 hours per day	-0.024	4,326	0.746
	(0.017)		
No of minutes spent on work	16.449***	4,326	518.388
	(6.281)		
Men's A-WEAI indicators			
Has some input in decisions/feels like can make decisions in at least 1 domain	-0.000	3,788	0.962
	(0.007)		

No of ag domains resp has some input in decision making	0.017	3,788	5.528
	(0.098)		
Solely/jointly owns at least 2 small or 1 large asset, incl land	-0.007*	3,788	0.998
	(0.004)		
No of assets that are solely/jointly owned by resp	-0.184**	3,788	6.868
	(0.078)		
Has some input in income dec or feels like can make decisions	0.003	3,788	0.973
	(0.006)		
No of domains resp has some input in income decisions/feels can make decision	0.261***	3,788	4.835
	(0.070)		
Jointly makes at least 1 decision regarding at least 1 source of credit	0.007	3,788	0.762
	(0.016)		
No. of accessible sources where solely/jointly participated in credit decisions	0.047	3,788	1.140
	(0.036)		
Resp is part of at least 1 group	0.123***	3,788	0.260
	(0.019)		
No of groups where resp is an active member	0.201***	3,788	0.314
	(0.027)		
Resp didn't work more than 10.5 hours per day	-0.010	3,788	0.660

	(0.019)		
No of minutes spent on work	20.025**	3,788	519.155
	(8.117)		
A-WEAI composite indicators			
Standardized weighted women's A-WEAI index - all indicators	0.086**	4,326	-0.000
	(0.035)		
Woman's five domains of empowerment (5DE) score	0.018***	4,326	0.876
	(0.006)		
Woman is empowered	0.055***	4,326	0.644
	(0.017)		
Standardized weighted men's A-WEAI index - all indicators	0.122***	3,788	-0.000
	(0.040)		
Man's five domains of empowerment (5DE) score	0.018**	3,788	0.840
	(0.008)		
Man is empowered	0.050**	3,788	0.571
	(0.020)		
Gap between man's and woman's 5DE scores within household	0.004	3,651	0.029
	(0.004)		

Household reached gender parity	0.033**	4,326	0.760
	(0.015)		

Notes: Estimates from the common support sample are selected by the matching model. Stars denote statistical significance as follows: * $p<0.10$; ** $p<0.05$; *** $p<0.01$

Appendix Table B.9: Impact estimates for child outcomes - Remove coastal districts

	Treatment effect (SE) from NNM	N	Comparison mean
Under 5 children: Length/height-for-age Z-score	-0.025	1,796	-1.493
	(0.080)		
Under 5 children: Stunting (HAZ < -2)	0.011	1,796	0.341
	(0.032)		
Under 2 children: Length/height-for-age Z-score	0.041	685	-1.312
	(0.141)		
Under 2 children: Stunting (HAZ < -2)	-0.075	685	0.302
	(0.050)		
IYCF practice: early initiation of breastfeeding for child 0-23mons	-0.050	692	0.772
	(0.049)		
IYCF practice: continued breastfeeding at 1 year for child 12-23mons	0.002	320	0.943
	(0.027)		
IYCF practice: Minimum dietary diversity (>=5 out of 8 food groups) for child 6-23mons	0.020	521	0.282

	(0.059)		
IYCF practice: Consumption of iron-rich food for child 6-23mons	0.198***	521	0.380
	(0.065)		

Notes: Estimates from the common support sample are selected by the matching model. Stars denote statistical significance as follows: * $p<0.10$; ** $p<0.05$; *** $p<0.01$

Appendix Table B.10: Impact estimates for agricultural production outcomes - Remove coastal districts

	Treatment effect (SE) from NNM	N	Comparison mean
Standardized weighted agricultural production index	-0.193***	2,805	-0.000
	(0.055)		
Planted HYV rice - all HHs	-0.032*	4,804	0.513
	(0.017)		
Planted HYV rice - all farmers	-0.099***	2,805	0.905
	(0.017)		
Planted HYV rice: TAus variety - all farmers	0.035**	2,805	0.053
	(0.014)		
Planted HYV rice: TAman variety - all farmers	-0.134***	2,805	0.644
	(0.022)		
Planted HYV rice: Boro variety - all farmers	-0.213***	2,805	0.744
	(0.022)		

Percentage of all rice out of total crop area - all farmers	-0.165***	2,805	0.817
	(0.016)		
Percentage of HYV rice out of total crop area - all farmers	-0.216***	2,805	0.785
	(0.017)		
Rice production/ha in metric ton - rice farmers	-0.026	2,551	3.169
	(0.056)		
HYV TAus rice production/ha in metric ton - HYV TAus rice farmers	-0.470	126	2.418
	(0.315)		
HYV TAman rice production/ha in metric ton - HYV TAman rice farmers	0.153**	1,717	2.731
	(0.063)		
HYV Boro rice production/ha in metric ton - HYV Boro rice farmers	0.486***	1,894	3.693
	(0.064)		
Rice variety production/ha in metric ton excluding local varieties	0.213***	2,445	3.251
	(0.057)		
Number of crops grown - all farmers	0.536***	2,805	1.663
	(0.067)		

Notes: Estimates from the common support sample are selected by the matching model. Stars denote statistical significance as follows: * $p<0.10$; ** $p<0.05$; *** $p<0.01$

Appendix Table B.11: Summary statistics of covariates of outside and inside ZOI samples at district level

	District inside ZOI			District outside ZOI			P-value	Normalized difference
	Mean	SD	N	Mean	SD	N		
Number of farming households	8,317.550	3,063.290	20	14,455.727	13,080.020	44	0.043	-0.646
Average operated land (acres)	1.308	0.226	20	1.387	0.637	44	0.597	-0.163
Average owned land (acres)	1.367	0.195	20	1.452	0.659	44	0.577	-0.174
Average number of family members in agriculture	2.819	0.138	20	3.008	0.250	44	0.003	-0.933
Average number of machines owned	1.332	0.189	20	1.354	0.138	44	0.598	-0.134
Poverty headcount	36.285	11.093	20	30.430	12.160	44	0.072	0.503

Appendix Table B.12: Impact estimates for poverty and food security outcomes - Restricted to comparable districts

	Treatment effect (SE) from NNM	N	Comparison mean
Poverty and food security - Districts in common support (42)			
Standardized weighted poverty index	0.001	2,926	0.000
	(0.040)		
Poor HH: daily per cap consumption expenditure < $1.25 PPP	0.011	2,926	0.246
	(0.018)		
Poverty gap where poverty line = $1.25 PPP	-0.154	2,926	5.032

	(0.453)		
HH Food Insecurity Experience Scale (raw score: 0-8)	-0.299***	2,926	1.705
	(0.082)		
Any Food Insecurity (FIES score >=1)	-0.132***	2,926	0.557
	(0.020)		
Mild Food Insecurity (1<=FIES<=3)	-0.146***	2,926	0.385
	(0.020)		
Moderate Food Insecurity (4<=FIES<=6)	0.015	2,926	0.148
	(0.015)		
Severe Food Insecurity (FIES>=7)	-0.002	2,926	0.024
	(0.008)		

Notes: Estimates from the common support sample are selected by the matching model at district level and at household level. The standardized weighted poverty index is constructed from (i) a dummy variable for being a poor household based on daily per capita consumption less than $1.25 PPP and (ii) poverty gap where poverty line at $1.25 PPP. The standardized weighted index outcome is aggregated following the procedure in Anderson (2008) such that the comparison group has mean of 0 and standard deviation of 1. Sharpened q-values in column 4 follow Benjamini et al. (2006) to control the false discovery rate. Stars denote statistical significance as follows: * $p<0.10$; ** $p<0.05$; *** $p<0.01$