The Journal of Nutrition
**Community and International Nutrition**

# Two Food-Assisted Maternal and Child Health Nutrition Programs Helped Mitigate the Impact of Economic Hardship on Child Stunting in Haiti[1,2]

Shannon Donegan,[3] John A. Maluccio,[3]* Caitlin K. Myers,[3] Purnima Menon,[4] Marie T. Ruel,[4] and Jean-Pierre Habicht[5]

[3]Department of Economics, Middlebury College, Middlebury, VT 05753; [4]Poverty, Health, and Nutrition Division, International Food Policy Research Institute, Washington, DC 20006; and [5]Division of Nutritional Sciences, Cornell University, Ithaca, NY 14853

## Abstract

Rigorous evaluations of food-assisted maternal and child health and nutrition programs are stymied by the ethics of randomizing recipients to a control treatment. Using nonexperimental matching methods, we evaluated the effect of 2 such programs on child linear growth in Haiti. The 2 well-implemented programs offered the same services (food assistance, behavior change communication, and preventive health services) to pregnant and lactating women and young children. They differed in that one (the preventive program) used blanket targeting of all children 6–23 mo, whereas the other (the recuperative program) targeted underweight (weight-for-age Z score $< -2$) children 6–59 mo, as traditionally done. We estimated program effects on height-for-age Z scores (HAZ) and stunting (HAZ $< -2$) by comparing outcomes of children in program areas with matched children from comparable populations in the Haiti Demographic and Health Survey. Children 12–41 mo in the preventive and recuperative program areas had lower prevalence of stunting than those in the matched control group [16 percentage points (pp) lower in preventive and 11 pp in recuperative]. Children in the 2 program areas also were more likely than those in the matched control group to be breast-fed up to 24 mo (25 pp higher in preventive, 22 in recuperative) and children 12 mo and older were more likely to have received the recommended full schedule of vaccinations (32 pp higher in preventive, 31 in recuperative). Both programs improved targeted behaviors and protected child growth in a time of deteriorating economic circumstances. J. Nutr. 140: 1139–1145, 2010.

## Introduction

In recent years, maternal and child nutrition have become high priorities for large food aid donors and implementing agencies (1). The United States Agency for International Development (USAID),[6] e.g., spends nearly \$200 million/y on food-assisted maternal and child health and nutrition (MCHN) programs.

There are few rigorous evaluations of such programs, however, and therefore little is known about their overall effectiveness in reducing undernutrition in different operational contexts. For example, available assessments typically use simple before-after comparisons without an adequate control group (2,3). Such designs do not permit attributing changes in the outcomes of interest to the intervention, because they lack a valid “counterfactual” or group that captures what would have happened in the absence of the intervention. These designs can thus be misleading in situations where changes in undernutrition occur over the evaluation period for reasons other than the intervention itself. This was the situation in Haiti between 2000 and 2005 when undernutrition rose as a result of economic, climatic, and political shocks (4,5). Programs that were assessed during this period using simple before-after comparisons appeared to have had little or even negative impact on child nutrition (3).

In previous work, we compared the impact on child anthropometry of 2 food-assisted MCHN programs implemented in Haiti from 2002 to 2005 (6). We showed that the preventive model, which used blanket targeting of all children 6–23 mo of age, was more effective at reducing stunting, wasting, and underweight prevalence compared with the recuperative model that targeted underweight children (6–59 mo of age). Our evaluation was not prejudiced by the trend in undernutrition in Haiti at the time because we used a cluster randomized trial to

[1] Supported by the Food and Nutrition Technical Assistance Project through USAID; USAID-Haiti; World Vision, Haiti; Government of Germany; and the World Food Program.

[2] Author disclosures: S. Donegan, J. A. Maluccio, C. K. Myers, P. Menon, M. T. Ruel, and J-P. Habicht, no conflicts of interest.

[6] Abbreviations used: ADP, Area Development Program; BCC, behavior change and communication; HDHS, Haiti Demographic and Health Survey; HAZ, height-for-age Z score; MCHN, maternal and child health and nutrition; pp, percentage points; USAID, United States Agency for International Development; WAZ, weight-for-age Z score; WV, World Vision.

*To whom correspondence should be addressed. E-mail: john.maluccio@middlebury.edu.


Manuscript received August 7, 2009. Initial review completed October 22, 2009. Revision accepted March 20, 2010.
First published online April 14, 2010; doi:10.3945/jn.109.114272.

allocate groups of communities to the 2 program models and compared child nutrition outcomes after communities had been exposed to either program model for 3 y. Both program models offered a range of services targeted to pregnant and lactating women and young children (6,7) and as such each was likely to be beneficial. For ethical, logistical, and financial reasons, however, the evaluation did not include a control group where food assistance was not provided. Therefore, we could not estimate the absolute effect of either program model on community prevalence of undernutrition relative to communities where there was no program (6,8).

In this study, we assessed the total absolute effects of each of the 2 program models on child linear growth in Haiti using nonexperimental matching techniques that are less biased than simple comparisons with nonrandomized control areas (9). We matched children included in the 2005 postintervention survey of the randomized trial in the preventive and recuperative program areas to comparison children drawn from the Haiti Demographic and Health Survey (HDHS) implemented at the same time. The overall aim was to document whether the 2 food-assisted MCHN program models helped mitigate the potential negative effects on child linear growth of Haiti's economic downturn during the first 5 y of the new millennium.

## Methods

We compared outcomes for children in the program areas in 2005 with outcomes for children outside the program areas also measured in 2005. Separately for each program, we assessed the absolute difference in mean outcomes between children in each program area and matched comparison children.

*Data sources.* We used 2 data sources in the analysis: *1*) the 2005 International Food Policy Research Institute-Cornell-World Vision (WV) evaluation survey conducted as part of the cluster randomized trial from May to September 2005 (hereafter, evaluation survey) (6,7); and, for comparison, *2*) the 2005–2006 HDHS collected from October 2005 to June 2006 (5).

The evaluation survey was conducted as part of the cluster-randomized trial to compare the 2 program models (preventive and recuperative) of food-assisted MCHN programs implemented by WV-Haiti in the Central department.

WV had been operating in the region for approximately a decade prior to the 2002 launch of the food-assisted MCHN programs examined in this article, implementing its Area Development Programs (ADP). These were private sponsorship programs that provided schooling and health inputs directly to individual sponsored children as well as more general health and development programs to their communities. Also prior to the 2002 programs, WV worked as a subcontractor to Catholic Relief Services implementing a USAID Title II MCHN program similar to the recuperative program. Through these activities, WV developed a strong rapport with the rural communities in the Central Plateau and gained important logistical experience working in the region, thereby creating an effective community-based platform. The later programs were based on this platform, facilitating their high quality implementation. In the design of the original impact evaluation, communities were matched on whether they had WV ADP (6).

In 2001, WV-Haiti itself became a direct implementer of a USAID Title II MCHN program targeting a population of 600,000. This program was set up to compare the newly developed preventive model with the traditional recuperative model; staffing, supervision, and all other programmatic operations were the same for both program models.

Both the preventive and recuperative program models offered a range of services described in detail elsewhere (6,7). In summary, the interventions included behavior change and communication (BCC) counseling at rally posts and mothers' clubs, and the provision of monthly food rations, both within a larger MCHN program that included food rations for pregnant and lactating mothers, pre- and postnatal consultations, home visits for newborn infants and severely undernourished children under 5 y of age, and the provision of vaccination, vitamin A supplementation, oral rehydration salts, and antihelminth drugs to all children under 5 y of age. Receipt of the food rations was conditional upon monthly participation in the rally posts and mothers' clubs. Food distribution and other components of the programs began in August 2002, except for the BCC package, which was implemented in May 2003.

The key differences between the preventive and the recuperative models were: *1*) the child eligibility criterion [age 6–23 mo or weight-for-age Z score (WAZ) $< -3$ for children 24–59 mo of age in preventive; WAZ $< -2$ for children 6–59 mo of age in recuperative]; *2*) the duration of child eligibility to receive the food (6–23 mo of age, up to 18 mo, in preventive; up to 9 mo for recuperative); and *3*) the content and focus, timing, sequencing, and total number of BCC sessions at the mothers' clubs. The BCC sessions in the preventive model focused on child feeding and care to prevent undernutrition and a precise schedule of 14 lessons was established to ensure that delivery of the age-specific counseling messages reached caregivers at the time they would be most beneficial. For recuperative, the BCC sessions focused on topics of relevance to undernourished children, such as causes of undernutrition, nutritious recipes, feeding during illness, and hygiene in food preparation, handling, and storage. All other aspects of the comprehensive MCHN program were the same for both program models.

The evaluation survey collected data on children (and their families) living in communities in the Central department where the 2 programs had been operating for ~3 y, regardless of whether they had actually received the intervention, as is necessary for an intent-to-treat evaluation (10). The survey was conducted in 20 areas (10 randomly allocated to the preventive model and 10 to the recuperative model) in 3 communes of the rural Central department that had not had WV food assistance programs prior to 2002 but in some instances had an ADP. Households were randomly selected from the set of all households with at least 1 child 12–41 mo of age (referred to as the index child) using as the sample frame census data collected by the research team just prior to the evaluation survey. These children were all born after the programs began and the youngest of them had been exposed to one or the other of the 2 programs for at least a full year by the time of the evaluation survey. The evaluation survey questionnaire, based on the HDHS questionnaire, was administered to the mother of the index child and collected data on household and maternal characteristics. Anthropometric measurements (height and weight) were taken on the index children and their caregivers.

The evaluation survey included 1500 index children: 748 in the preventive and 752 in the recuperative group. We excluded 22 children with missing anthropometric measures or incomplete characteristics on which we matched them to the HDHS (as described below). After constructing Z scores for height-for-age (HAZ), we also excluded 2 observations with outlier values based on WHO criteria (11). The final evaluation survey sample included 731 children in the preventive group and 745 in the recuperative group.

The evaluation survey data were compared with data from the 2005–2006 HDHS, a nationally representative household survey, conducted in all 10 departments of Haiti (5). The HDHS used a 2-stage sampling scheme, with selection at the first stage from clusters with probability proportional to size followed by a second stage of random selection of households within each cluster (5,12). A household questionnaire included many of the same household and maternal characteristics as the evaluation survey and anthropometric measurements were taken on children under 5 y of age.

To ensure that the children from the HDHS who were used for comparison lived in conditions similar to those of children in the evaluation survey, we retained from the HDHS only those children 12–41 mo of age living in rural areas of the country surveyed from October to December 2005. Children living in urban areas were not comparable to those in the evaluation survey, because the prevalence and severity of undernutrition and food insecurity, as well as the access to health services, differed markedly between urban and rural areas. In addition, only children measured in 2005 would have experienced the changes in food security and environmental conditions caused by the natural disasters and civil disruptions at the same ages as the intervention children. This was particularly important, because the effects of such shocks might have had different impacts at different ages. The effect on anthropometry of the few months' difference between the surveys should have been minor for HAZ, a cumulative indicator of past nutrition, but is likely to have been more substantial for WAZ and weight-for-height Z scores, because these latter 2 indicators are more responsive to short-term fluctuations in the environment such as seasonal changes. For this reason, WAZ and weight-for-height Z score were not analyzed. As with the evaluation survey, we excluded children with missing anthropometric measures or incomplete characteristics on which we matched or with outlier values for Z scores according to the WHO criteria (11). The final HDHS sample available for potential matches included 355 children.

Direct evidence that both programs were implemented equally well is provided elsewhere (7,13,14). To complement and confirm these findings with indirect evidence on effective implementation, we first examined 3 binary indicators related to the impact of the program BCC and health care service components: *1*) whether a child older than 12 mo of age was breast-fed up to 12 mo of age; *2*) whether a child older than 24 mo of age was breast-fed up to 24 mo of age, the WHO-recommended duration for breast-feeding (15); and *3*) whether children 12 mo of age or older were fully vaccinated as reported on a vaccine card or by the mother when the card was not available. The vaccines included the vaccine against tuberculosis, Bacille Calmette-Guérin, the measles, mumps, and rubella vaccine, the 3rd dose of oral polio vaccine, and the 3rd dose of the diphtheria-pertussis-tetanus vaccine.

The main outcomes examined to assess linear growth were mean HAZ scores and the prevalence of stunting, defined as HAZ $< -2$.

*Statistical analysis.* To estimate program effects, an appropriate comparison group was necessary. For a comparison group to be valid, it should be comparable enough to the intervention group so that differences between the 2 groups can be attributed to the treatment. This comparability can be achieved by finding a group that did not have the program but had levels of the determinants of the outcomes (e.g. HAZ) similar to those of the intervention group. These determinants, which are confounders of any real program effect, can be controlled for by matching each child in the intervention group with a comparison child who is exposed to all of the same determinants of the outcome under consideration, except the treatment. This would yield both maximum comparability and the greatest number of comparison children.

An approximation to this ideal is now computationally feasible using nonexperimental matching methods with multiple comparison children which, with recent advancements (16), has become an increasingly common way to estimate program impacts for antipoverty and related programs (9,16–20). We used this method to estimate the mean difference in outcomes between treatment children in communities that received the WV-Haiti programs (treated) and similar matched children from the HDHS survey (untreated). This method would ensure complete comparability if all of the confounding determinants of anthropometry were known and perfectly measured. Because this was not possible, we must assume that, conditional on matching individuals on a vector of measurable characteristics, untreated individuals had the same mean outcomes that the treated individuals would have had if they had not received the program. This is known variously as the assumption of selection on observables, the conditional independence assumption, or unconfoundedness (16,21).

Direct comparisons of nonexperimental matching methods with experimental estimators based on randomized designs have shown that the nonexperimental methods are more likely to satisfy the unconfoundedness assumption and therefore provide reliable, low-bias estimates of program effects when the data include covariates associated with the outcomes being examined as well as with treatment status, the data sources used are similar (e.g. the same questions and survey methods), and the local conditions are similar (22–25). The first 2 conditions are theoretically necessary and substantiated empirically in the literature, and they are met in our study. The 3rd condition is an empirical finding in the literature meaning that more unmeasured confounders are present when the geographic areas are different. We describe next the steps we implemented to fulfill the 3rd condition. Finally, we assessed the sensitivity of the results to the choice of methodology and matching variables used, because estimates may be sensitive to these factors (20). These steps were all aimed at increasing the plausibility of unconfoundedness.

First, guided by theoretical considerations (and subject to variables available in each of the 2 data sets), we determined the potential set of variables upon which the matches could be based (**Table 1**). The variables chosen should have predictive power for the outcomes under study as well as for treatment status (20,26–28). Because we used data collected in the intervention areas after the programs began, it was also important to choose variables that were unlikely themselves to have been influenced substantially by the programs. Second, we constructed a propensity score for each child to guide the final selection of the exact form of the matching variables (including polynomial transformations and interactions). The propensity score model was the probit of living in a treatment area as explained by the variables shown in Table 1. We calculated the probit so that the distributions of the variables in Table 1 (and any resulting transformations) for each quantile of the propensity score were statistically the same across program and HDHS children, a procedure referred to as balancing (20). We then used these balance variables or their transformations to match each child in the preventive program area (and, separately, each child in the recuperative program area) to the 2 most similar children ("nearest neighbors") from the HDHS. The propensity score itself was not used for matching. We matched children exactly on a 3-mo age range (e.g. 12 to <15 mo, 15 to <18 mo, and so on up to 39 to <42 mo) and using nearest neighbor matching on the other variables (and their transformations) (27,29).

The estimated average treatment effect is the difference in mean outcomes for the treated compared with the matched groups where, as a result of allowing multiple matches, the matched group may include repeated observations on some children. To calculate the SE and corresponding confidence intervals and *t*-tests, we implemented a heteroskedasticity robust variance estimator that is standard for this nearest neighbor matching technique (29). This also entailed a matching procedure to match observations to a nearest neighbor but this time from within the same treatment category instead of across treatment groups, as discussed above. Then the sample variance within each of these matched pairs was calculated and used to estimate the conditional variance of the sample mean outcome for each group. The variance of the average treatment effect estimator is the sum of the estimated variances of the samples means for the treated and matched groups.

We also undertook several sensitivity analyses using: *1*) alternative matching methods such as nearest neighbor matching with between 1 and 5 neighbors, as well as kernel matching (16); *2*) alternative sets of matching variables; and *3*) alternative subsamples of the HDHS as the source of comparison observations. All analyses were carried out using Stata version 11 (30).

We set significance at a 2-tailed $P < 0.05$.

*Ethical review.* Approval for the data collection and related analyses for the evaluation study was obtained from the Cornell University Committee on Human Subjects. The secondary data analysis presented in this article was exempted from ethical review, because we used de-identified data from the International Food Policy Research Institute-Cornell-WV evaluation survey data set and the 2005–2006 HDHS data set, available for use with permission from Measure DHS.

## Results

Of the 355 children 12–41 mo of age in the HDHS subsample from rural areas measured in 2005 (i.e. the potential matches), 290 were matched to at least 1 child in the preventive program and 300 to at least 1 child in the recuperative program (Table 1). When combined, there were 325 unique children from HDHS matched to 1 or more children in 1 or both of the treatment groups. Children from the 2 treatment groups were very similar to each other, which is unsurprising given that they were from the same region and the type of treatment was randomly

**TABLE 1** Summary statistics by sample[1]

| | Program | | HDHS | |
|---|---|---|---|---|
| | Preventive | Recuperative | Matched to preventive | Matched to recuperative |
| Child's characteristics | | | | |
| Age, *mo* | 24.6 ± 8.3 | 26.0 ± 8.7 | 25.9 ± 8.5 | 26.2 ± 8.7 |
| Male, *%* | 48.6 | 50.5 | 47.2 | 47.7 |
| Mother's characteristics | | | | |
| Age, *y* | 29.7 ± 6.8 | 29.3 ± 6.7 | 30.4 ± 7.1 | 30.4 ± 7.0 |
| Height, *cm* | 158.4 ± 5.9 | 157.8 ± 6.8 | 158.3 ± 6.0 | 158.0 ± 5.9 |
| Education, *y* | 2.0 ± 2.7 | 1.9 ± 2.7 | 2.0 ± 2.7 | 2.0 ± 2.8 |
| Completed primary education, *%* | 45.0 | 43.9 | 42.1 | 40.3 |
| Completed secondary education, *%* | 8.1 | 7.9 | 8.3 | 8.3 |
| Literate, *%* | 39.3 | 39.9 | 33.8 | 34.3 |
| Father's characteristics | | | | |
| Present in household, *%* | 92.9 | 91.5 | 93.4 | 94.0 |
| Education, *y* | 2.8 ± 3.4 | 2.7 ± 3.3 | 3.1 ± 3.4 | 3.1 ± 3.4 |
| Completed primary education, *%* | 46.9 | 46.4 | 47.2 | 46.0 |
| Completed secondary education, *%* | 15.6 | 14.4 | 15.5 | 15.7 |
| Farming is main occupation, *%* | 86.9 | 86.3 | 72.1 | 72.0 |
| Assets and dwelling | | | | |
| Radio, *%* | 40.5 | 33.0 | 40.7 | 40.7 |
| Television, *%* | 0.0 | 0.7 | 0.0 | 1.0 |
| Bicycle, *%* | 11.5 | 12.1 | 6.9 | 8.0 |
| Motorcycle, *%* | 0.5 | 1.1 | 0.0 | 1.0 |
| Earth floor, *%* | 94.8 | 95.0 | 85.5 | 82.7 |
| Thatch roof, *%* | 50.2 | 51.0 | 26.9 | 26.7 |
| Toilet, *%* | 38.9 | 36.2 | 36.6 | 36.3 |
| Protected water source, *%* | 36.0 | 37.9 | 30.0 | 31.3 |
| Time to collect water, *min* | 37.3 ± 29.8 | 39.9 ± 33.9 | 28.3 ± 30.3 | 30.5 ± 34.2 |
| Household composition | | | | |
| Children, *n* | 3.7 ± 1.6 | 3.7 ± 1.7 | 3.5 ± 1.7 | 3.6 ± 1.7 |
| Adults, *n* | 2.9 ± 1.3 | 2.8 ± 1.4 | 2.8 ± 1.3 | 2.8 ± 1.4 |
| Observations, *n* | 731 | 745 | 290 | 300 |

[1] Values are mean ± SD or percent.

JN THE JOURNAL OF NUTRITION

assigned (6). Treatment children in each of the 2 groups also were quite similar to those children in the HDHS sample for each analysis. Children from the program samples, however, were significantly more likely to have fathers who reported farming as their primary occupation and to live in households that were slightly worse off as measured by asset ownership than were the HDHS children.

The mean differences between each program child and his/her 2 matched children from the HDHS indicated that children living in preventive and recuperative areas had substantially and significantly higher reported rates of 2 important outcomes promoted by the programs: breast-feeding and full vaccination coverage (**Table 2**). The prevalence of breast-feeding up to 12 mo of age was >97% in both program areas and was 12.7 percentage points (pp) higher for children in preventive areas than for matched children in the HDHS, and 12.5 pp higher for children in recuperative areas. The prevalence of breast-feeding until the WHO recommendation of 24 mo (15) was 25.3 pp higher for children in preventive areas and 21.7 pp higher for children in recuperative areas. Full vaccination coverage for children 12 mo of age and older was >75% in both program areas and was 31.7 pp higher for children in preventive areas and 30.7 pp for children in recuperative areas.

After ~3 y of operations, children 12–41 mo of age living in preventive or recuperative program areas also had generally greater linear growth than matched children from the HDHS (**Table 3**). Children from preventive areas had higher mean HAZ scores (+0.341) than matched children from the HDHS. These differences in mean Z scores translated into lower prevalence rates for stunting, which was 15.9 pp lower in the preventive areas. Children from recuperative areas also had higher mean HAZ scores than matched children from the HDHS (+0.183; $P$ = 0.08), although this difference was not as large as for the preventive group and was not significant. The difference in mean Z scores corresponded to a 10.7-pp lower prevalence rate of stunting for children in the recuperative areas.

## Discussion

This study shows that both the preventive and recuperative models of a food-assisted MCHN program led to higher mean HAZ scores and a lower prevalence of stunting among children living in communities exposed to the programs compared with children from rural Haiti who were not exposed. Moreover, the estimated effects were substantial, suggesting, e.g., reductions in the prevalence of stunting of ~3 pp/y for the recuperative program, similar to effects previously reported in the literature using less robust nonexperimental designs (2), and 5 pp/y for the preventive program, a program type for which there are no previous results in the literature.

**TABLE 2** Nearest-neighbor matching estimates of differences in breast-feeding and vaccinations for all HDHS[1]

| | Preventive | | | Recuperative | | |
|---|---|---|---|---|---|---|
| | Preventive | Matched HDHS (all) | Difference | Recuperative | Matched HDHS (all) | Difference |
| | | | % | | | |
| Breast-fed to 12 mo | 97.3 | 84.7 | 12.7 (5.9, 19.5) | 97.9 | 85.5 | 12.5 (6.8, 18.1) |
| Breast-fed to 24 mo | 45.7 | 20.5 | 25.3 (14.7, 35.9) | 38.5 | 21.7 | 16.9 (6.8, 27.0) |
| Fully vaccinated[2] | 77.3 | 45.7 | 31.7 (18.0, 45.4) | 75.5 | 44.9 | 30.7 (18.6, 42.9) |

[1] Values are mean percentages with the 95% CI for the differences. Nearest-neighbor matching results using 2 nearest neighbors and requiring exact matches on 3-mo range age dummies.
[2] The vaccines include the vaccine against tuberculosis, Bacille Calmette-Guérin, the measles, mumps, and rubella vaccine, the third dose of oral polio vaccine, and the third dose of the diphtheria-pertussis-tetanus vaccine.

These large effects were realized in one of the poorest regions of the country, with the highest prevalence of stunting, and during a period when Haiti was undergoing economic, climatic, and political hardship (31). Purposively, WV-Haiti targeted only relatively poor areas of the country. HDHS data from 2000 and 2005 indicate that the Central department where the 2 programs operated was consistently one of the poorest, as measured by asset ownership or a wealth index. This relative poverty, within the poorest country in Latin America, also was reflected in the region's nutritional indicators. In both 2000 and 2005, the Central department had the highest prevalence of stunting in the country for children <5 y of age. Between 2000 and 2005, the prevalence of undernutrition soared in the region; stunting increased by 5 pp (37.9 to 42.8%) and there was a doubling of the prevalence of underweight (12.2 to 24.9%) and wasting (4.5 to 8.4%). As a result, the WV-Haiti programs, which yielded equal or lower prevalence of stunting in the program areas compared with baseline measurements taken for the randomized trial in 2002, effectively operated as safety nets for their beneficiaries during a downturn, mitigating the potential deterioration.

Several aspects of the results point to the plausibility of our findings (32). First, the examination of the effects on outcomes other than anthropometry that were targeted by both programs demonstrates that key aspects of the programs were being implemented and were highly effective. For full vaccination and breast-feeding up to 12 mo, the effects were similar in size across the 2 program models, consistent with there being no design or implementation differences related to these outcomes across the programs. Breast-feeding was emphasized in the mothers' clubs and vaccination was programmed to be done during the rally post visits in the first 6 mo of life. Both program models targeted lactating women (up to 6 mo postpartum) and their participation was approximately equal across the 2 models (7). The effect on the prevalence of breast-feeding up to 24 mo of age was higher in the preventive group, however, possibly because all mothers with children aged 6–23 mo had to participate in the mothers' clubs to receive the food rations, whereas in the recuperative program, only the mothers of underweight children were eligible for food rations and therefore required to attend mothers' clubs.

Second, the differences between the estimated absolute effects for the 2 programs reported here are consistent with, and nearly identical to, the experimental results reported previously (6), where it was shown that the preventive model was more effective than the recuperative model. Because it was more rigorous in the present analysis to match children separately for the preventive and recuperative programs, we did not match exactly the same children for each separate assessment. This means that the difference between the absolute effect for the preventive and recuperative programs reported above need not have been the same as the difference based on the randomized trial.

Finally, in addition to the matching methods presented here (29), we considered alternative matching techniques and sets of matching variables (20), all of which yielded similar results. The alternative methods considered included varying the number of neighbors used in the match from 1 to 5, nearest neighbor with bias correction (29), matching exactly on both 3-mo age groups and gender of the child, and Gaussian kernel matching, based directly on the estimated propensity scores (20,23). Our findings, therefore, are robust to the matching techniques used. We also examined results excluded from the variables used for matching the asset and household composition variables (Table 1), given the potential concern that the programs may have affected these confounders directly. The results were the same. This is evidence that the selection on unobservables assumption is maintained, because altering the set of observables characteristics on which we match, themselves likely to be correlated with important unobservables, did not alter the results. Finally, for matching variables that showed some imbalance between the treatment and matched samples (father's occupation and the dwelling's roofing and flooring materials), we reestimated the matching models requiring exact matches on those variables in addition to the 3-mo age group variable. Results were unchanged, indicating the findings are robust to more exact matching on these most likely confounding variables.

**TABLE 3** Nearest-neighbor matching estimates of differences in HAZ and stunting for all HDHS[1]

| | Preventive | | | Recuperative | | |
|---|---|---|---|---|---|---|
| | Preventive | Matched HDHS (all) | Difference | Recuperative | Matched HDHS (all) | Difference |
| HAZ | −1.519 | −1.859 | 0.341 (0.104, 0.577) | −1.672 | −1.855 | 0.183 (−0.022, 0.388) |
| Stunted, % | 33.3 | 49.1 | −15.9 (−25.7, −6.0) | 38.8 | 49.5 | −10.7 (−20.0, −1.5) |

[1] Values are means with the 95% CI for the differences. Nearest-neighbor matching results using 2 nearest neighbors and requiring exact matches on 3-mo range age dummies.

**TABLE 4** Nearest-neighbor matching estimates of differences in HAZ and stunting for alternative HDHS samples[1]

| | HDHS (all) | HDHS (excluding Central department) | HDHS (excluding regions with food programs) |
|---|---|---|---|
| Preventive-matched HDHS | | | |
| HAZ | 0.341 (0.104, 0.577) | 0.394 (0.137, 0.650) | 0.787 (0.388, 1.185) |
| Stunted, % | −15.9 (−25.7, −6.0) | −15.0 ( −24.4, −5.6) | −27.0 (−42.2, −11.7) |
| Recuperative-matched HDHS | | | |
| HAZ | 0.183 (−0.022, 0.388) | 0.234 (0.015, 0.454) | 0.602 (0.261, 0.943) |
| Stunted, % | −10.7 (−20.0, −1.5) | −9.8 (−18.7, −0.8) | −19.5 (−33.8, −5.2) |

[1] Values are mean differences (95% CI). Nearest-neighbor matching results using 2 nearest neighbors and requiring exact matches on 3-mo range age dummies.

There are some limitations to our study. WV-Haiti targeted only relatively poor areas of the country. As a result, there remains a concern that matched children from the HDHS national rural sample may still have been relatively better off than those in the evaluation survey, as suggested by the mean characteristics. A related concern is that food aid programs were common throughout rural Haiti during this period. For example, WV-Haiti was 1 of 4 organizations operating a USAID food-assisted MCHN program during this period, with the other organizations operating in portions of 3 of the 4 departments interviewed in HDHS in 2005 (3) and therefore included in the comparison group. Finally, we retained in the comparison sample households in the HDHS from the Central department, because they were likely to be geographically very similar to WV-Haiti program households, a key consideration for successful matching. These features raise the possibility that households selected for matching also were beneficiaries of a food aid program, including the WV-Haiti program itself. In areas of operation outside the evaluation survey communities, WV-Haiti employed the recuperative model. Together, the relative poverty of the Central department and the widespread nature of food aid in Haiti both suggest that comparisons made with the HDHS may lead to underestimates of WV-Haiti program effects, comparing treated children with children from slightly wealthier backgrounds or children in households also receiving food aid or both. Therefore, we examined the use of alternative HDHS comparison samples, excluding the Central department and, separately, drawing from departments without large food aid programs at the time (**Table 4**). Exclusion of the Central department led to similar and, in most cases, slightly larger differences between treatment and comparison groups, including the difference in HAZ for the recuperative model, which is now significant. Exclusion of the departments with large food programs revealed differences for the linear growth measures that were substantially larger. On balance, consideration of these alternative comparison samples led to differences in the estimated effects on HAZ in expected ways and program effects reported in Table 3 appear, if anything, to be conservative.

Rigorous evaluations of food-assisted MCHN programs are stymied by the ethics of randomizing recipients to a control treatment. As a result, little is known about their overall effectiveness in reducing undernutrition in different operational contexts. Using nearest neighbor matching methods, we estimated the absolute effects of 2 well-executed programs of this type implemented by WV in the Central department. Children living in both the preventive and recuperative model areas had greater mean HAZ and lower rates of stunting compared with matched control children. Communities exposed to the preventive program, previously shown to have been more effective than the recuperative program (6), had a prevalence of stunting 16 pp lower than the matched comparison. The difference in stunting prevalence for communities exposed to the recuperative program compared with the matched control group was smaller, 11 pp, but still substantial. In conclusion, both food-assisted MCHN programs were effective at tackling child undernutrition in this context characterized by economic hardship. Methodologically, the matching-based analytic approach used to create a control group in this study offers promise for other study settings where identifying or setting up a control group is challenging.


### Acknowledgments

This research began as S. Donegan's undergraduate thesis project at Middlebury College and, along with J.A. Maluccio and C.K. Myers, she designed the analysis and wrote the manuscript. M.T. Ruel was principal investigator of the randomized trial component of the study, which was designed with P. Menon and J-P. Habicht, among others. All authors critically analyzed and interpreted the data, and read and approved the final version of the article.


## Literature Cited


1. Barrett CB, Maxwell DG. Food aid after fifty years: recasting its role. New York: Routledge; 2005.
2. Swindale A, Deitchler M, Cogill B, Marchione T. The impact of Title II maternal and child health and nutrition programs on the nutritional status of children. Washington, DC: Food and Nutrition Technical Assistance Project Occasional Paper No. 4. Academy for Educational Development; 2004.
3. Merritt CG, Frankel NL, Hopstock P, Schlossman N, Vilgrain C, Charles G, Li X, Young M. Evaluation of the PL 480 Title II Program: USAID/Haiti Development Assistance Program, Report delivered to USAID, 22 January 2007. Arlington (VA): Development Associates; 2007.
4. Cayemittes M, Placide MF, Barrère B, Mariko S, Sévère B. Enquête mortalité, morbidité et utilisation des services, Haïti 2000. DHS. Calverton (MD): Ministère de la Santé Publique et de la Population, Institut Haïtien de l'Enfance and ORC Macro; 2001.
5. Cayemittes M, Placide MF, Mariko S, Barrère B, Sévère B, Alexandre C. Enquête Mortalité, Morbidité, et Utilisation des Services, Haïti 2005–2006. Calverton (MD): Ministère de la Santé Publique et de la Population, Institut Haïtien de l'Enfance and ORC Macro; 2007.
6. Ruel MT, Menon P, Arimond M, Bergeron G, Habicht JP, Hankebo B, Loechl C, Maluccio JA, Michaud L, et al. Age-based preventive targeting of food assistance and behaviour change and communication for reduction of childhood undernutrition in Haiti: a cluster randomised trial. Lancet. 2008;371:588–95.
7. Menon P, Ruel MT, with contributions from Arimond M, Habicht JP, Hankebo B, Loechl C, Maluccio J, Mbuya M, Michaud L, Pelto G. Prevention is better than cure. Final report of the evaluation: Prevention or Cure? Comparing preventive and recuperative approaches to

targeting maternal and child health and nutrition programs in rural Haiti. Report submitted to the Food and Technical Assistance (FANTA) Project, July 31, 2007. Washington DC, International Food Policy Research Institute, 2007.
8. White H. Tackling childhood undernutrition. Lancet. 2008;371:539–41.
9. Ravallion M. Evaluation of antipoverty programs. In: Schultz TP, Strauss J, editors. Handbook of development economics. 4th vol. New York: Elsevier; 2008.
10. Greenland S. Introduction to regression modeling. In: Rothman KJ, Greenland S, editors. Philadelphia: Lippencott; 1998. p. 425.
11. WHO. Anthro for personal computers, version 2: software for assessing growth and development of the world's children. Geneva: WHO; 2007.
12. Macro International. Sampling manual. DHS-III basic documentation series. Calverton (MD): Macro International; 1996.
13. Loechl C, Ruel MT, Pelto G, Menon P. An operations evaluation of World Vision's Integrated Health and Nutrition Program in Central Plateau, Haiti. Submitted to the Food and Technical Assistance (FANTA) Project, March 2004. Washington (DC): International Food Policy Research Institute; 2004.
14. Menon P, Mbuya M, Habicht JP, Pelto G, Loechl CU, Ruel MT. Assessing supervisory and motivational factors in the context of a program evaluation in rural Haiti. J Nutr. 2008;138:634–7.
15. WHO. Global strategy for infant and young child feeding. Geneva: WHO; 2003.
16. Imbens GW, Wooldridge JM. Recent developments in the econometrics of program evaluation. J Econ Lit. 2009;47:5–86.
17. Jalan J, Ravallion M. Estimating the benefit incidence of an antipoverty program using propensity score matching. J Bus Econ Stat. 2003;21:19–35.
18. Levine D, Painter G. The schooling costs of teenage out-of-wedlock childbearing: analysis with a within-school propensity-score-matching estimator. Rev Econ Stat. 2003;85:884–900.
19. Gilligan DO, Hoddinott J. Is there persistence in the impact of emergency food aid? Evidence on consumption, food security, and assets in rural Ethiopia. Am J Agric Econ. 2007;89:225–42.
20. Todd P. Evaluating social programs with endogenous program placement and selection of the treated. In: Schultz TP, Strauss J, editors. Handbook of development economics. 4th vol. New York: Elsevier; 2008. p. 3847–94.
21. Rosenbaum P, Rubin D. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983;70:41–50.
22. Heckman J, Ichimura H, Todd P. Matching as an econometric evaluation estimator: evidence from evaluating a job training program. Rev Econ Stud. 1997;64:605–54.
23. Heckman J, Ichimura H, Todd P. Matching as an econometric evaluation estimator. Rev Econ Stud. 1998;65:261–94.
24. Heckman J, Ichimura H, Smith J, Todd P. Characterizing selection bias using experimental data. Econometrica. 1998;66:1017–89.
25. Díaz JJ, Handa S. An assessment of propensity score matching as a non-experimental impact estimator: evidence from Mexico's *PROGRESA* Program. J Hum Resour. 2006;41:319–45.
26. Abadie A, Imbens GW. Large sample properties of matching estimators for average treatment effects. Econometrica. 2006;74:235–67.
27. Heckman J, Navarro-Lozano S. Using matching, instrumental variables, and control functions to estimate economic choice models. Rev Econ Stat. 2004;86:30–57.
28. Smith J, Todd P. Does matching overcome LaLonde's critique of non-experimental estimators? J Econom. 2005;125:305–53.
29. Abadie A, Drukker D, Leber Herr J, Imbens GW. Implementing matching estimators for average treatment effects in Stata. Stata Journal. 2004;4:290–311.
30. StataCorp. Stata statistical software: release 11. College Station (TX): Stata Corporation; 2009.
31. International Crisis Group. Haiti's transition: hanging in the balance. Latin American/Caribbean Briefing No. 7, February 8, 2005. New York: 2005; International Crisis Group.
32. Habicht JP, Victora CG, Vaughan JP. Evaluation designs for adequacy, plausibility and probability of public health program performance and impact. Int J Epidemiol. 1999;28:10–8.