IEG WORLD BANK | IFC | MIGA
INDEPENDENT EVALUATION GROUP

What Can We Learn from Nutrition Impact Evaluations?

LESSONS FROM A REVIEW OF INTERVENTIONS TO REDUCE CHILD MALNUTRITION IN DEVELOPING COUNTRIES

The World Bank Group

WORKING FOR A WORLD FREE OF POVERTY

The World Bank Group consists of five institutions—the International Bank for Reconstruction and Development (IBRD), the International Finance Corporation (IFC), the International Development Association (IDA), the Multilateral Investment Guarantee Agency (MIGA), and the International Centre for the Settlement of Investment Disputes (ICSID). Its mission is to fight poverty for lasting results and to help people help themselves and their environment by providing resources, sharing knowledge, building capacity, and forging partnerships in the public and private sectors.

The Independent Evaluation Group

IMPROVING DEVELOPMENT RESULTS THROUGH EXCELLENCE IN EVALUATION

The Independent Evaluation Group (IEG) is an independent, three-part unit within the World Bank Group. IEG-World Bank is charged with evaluating the activities of the IBRD (The World Bank) and IDA, IEG-IFC focuses on assessment of IFC's work toward private sector development, and IEG-MIGA evaluates the contributions of MIGA guarantee projects and services. IEG reports directly to the Bank's Board of Directors through the Director-General, Evaluation.

The goals of evaluation are to learn from experience, to provide an objective basis for assessing the results of the Bank Group's work, and to provide accountability in the achievement of its objectives. It also improves Bank Group work by identifying and disseminating the lessons learned from experience and by framing recommendations drawn from evaluation findings.

IEG WORLD BANK | IFC | MIGA

INDEPENDENT EVALUATION GROUP

LESSONS FROM A REVIEW OF INTERVENTIONS TO REDUCE CHILD MALNUTRITION IN DEVELOPING COUNTRIES

What Can We Learn from Nutrition Impact Evaluations?

2010
The World Bank
Washington, D.C.

1818 H Street, N.W.
Washington, D.C. 20433
Telephone: 202-473-1000
Internet: www.worldbank.org
E-mail: feedback@worldbank.org

1 2 3 4 13 12 11 10

Cover: Girl being weighed during a routine checkup at the Santa Rosa de Lima clinic in Nueva Esperanza, Honduras, which specializes in health care for children. Photo by Alfred Srur, courtesy of the World Bank Photo Library.

ISBN-13: 978-0-8213-8406-0
e-ISBN-13: 978-0-8213-8407-7
DOI: 10.1596/978-0-8213-8406-0

Library of Congress Cataloging-in-Publication Data have been applied for.

World Bank InfoShop
E-mail: pic@worldbank.org
Telephone: 202-458-5454
Facsimile: 202-522-1500
Printed on Recycled Paper

Independent Evaluation Group
Communication, Strategy, and Learning
E-mail: eline@worldbank.org
Telephone: 202-458-4497
Facsimile: 202-522-3125

Printed on Recycled Paper

Table of Contents

Boxes

Figures

Tables

Abbreviations

BINP	Bangladesh Integrated Nutrition Project
BDH	*Bono de Desarrollo Humano*
CAC	*Componente de Atención a Crisis*
CCT	Conditional cash transfer
CENP	Community Empowerment and Nutrition Program, Vietnam
CNP	Community nutrition promoter
CNW	Community nutrition worker
DALY	Disability-adjusted life-year
DHS	Demographic and Health Survey
DID	Difference-in-difference
DIME	Development Impact Evaluation Initiative
FA	*Familias en Acción*
HAZ	Height-for-age z-score
HIV	Human immunodeficiency virus
HNP	Health, nutrition, and population
ICDS	Integrated Child Development Services
ICR	Implementation Completion and Results Report
IEG	Independent Evaluation Group
IMCI	Integrated Management of Childhood Illness
ITT	Intent-to-treat
IV	Instrumental variables
LBW	Low birthweight
MDG	Millennium Development Goal
NGO	Nongovernmental organization
PAD	Project Appraisal Document
PIDI	*Proyecto Integral de Desarrollo Infantil* (Integrated Child Development Project), Bolivia
PLW	Pregnant or lactating women
PRN	*Programme de Renforcement de la Nutrition* (Nutrition Enhancement Program), Senegal
PROGRESA	*Programa Nacional de Educación, Salud y Alimentación* (National Program for Education, Health and Nutrition), Mexico (now *Oportunidades*)
PSM	Propensity score matching
RCT	Randomized controlled trial
RPS	*Red de Protección Social*
SCF	Save the Children Federation
SD	Standard deviation
SEECALINE	*Projet de Surveillance et Education des Écoles et des Communautés en Matière d'Alimentation et de Nutrition Élargi* (Expanded Project for Monitoring and Education of Schools and Communities in Food and Nutrition), Madagascar
THR	Take-home rations
UCT	Unconditional cash transfer
UNICEF	United Nations Children's Fund
WAZ	Weight-for-age z-score
WHZ	Weight-for-height z-score

Acknowledgments

This report was prepared by Martha Ainsworth (Task Manager) and Alemayehu Ambel, with inputs from Ximena del Carpio, Gayle Martin, and Shampa Sinha, under the supervision of Monika Huppi. The report was edited by William Hurlbut and Linda Harteker, and administrative support was provided by Marie-Jeanne Ndiaye. The external reviewers for the report were Alessandra Marini and Agnes Quisumbing.

The authors wish to thank Jean-Pierre Habicht, Yi-Kyoung Lee, and Meera Shekar for their feedback and comments at various stages of preparation of the report, and Anupa Bhaumik of the Research Committee staff for facilitating access to the research proposals for several of the impact evaluations. Comments from the following individuals on previous drafts are much appreciated: Javier Baez, Arup Banerji, Hans-Martin Boehmer, Lynn Brown, Nils Fostvedt, Emauela Galasso, Ali Khadr, Nestor Ntungwanayo, Meera Shekar, J.P. Singh, Mark Sundberg, Marvin Taylor-Dormond, Denise Vaillancourt, Christine Wallich, and Andrew Warner.

We also thank the following researchers and task managers of World Bank projects with nutrition impact evaluations and developing country policy makers interviewed for chapter 3: Tahmeed Ahmed, Harold Alderman, Orazio Attanasio, Jere Behrman, Carla Bertoncino, Lynn Brown, Donald Bundy, Manuel Fernando Castro, Khadidiatou Dieng, Daniel Dulitzky, Emanuela Galasso, Marito Garcia, Michele Gragnolati, Nelson Gutierrez, Theresa Ho, Polly Jones, Elizabeth King, Kees Kostermans, John Newman, Rekha Menon, Menno Mulder-Sibanda, Rahman Mustafiz, John Mutumba, Peter Okwero, Norotiana Rakotomalala, Laura Rawlings, Claudia Rokx, Sandra Rosenhouse, Norbert Schady, Maryanne Sharp, Meera Shekar, and Howard White.

The opinions expressed in this report are those of the authors, who are responsible for the interpretations and any omissions.

Director-General, Evaluation: *Vinod Thomas*
Director, Independent Evaluation Group (IEG)–World Bank: *Cheryl Gray*
Manager, IEG Sector Evaluation: *Monika Huppi*
Task Manager: *Martha Ainsworth*

Executive Summary

High levels of child malnutrition in developing countries contribute to mortality and have long-term consequences for children's cognitive development and earnings in adulthood. Recent impact evaluations show that many interventions have had an impact on children's anthropometric outcomes (height, weight, and birthweight), but there is no simple answer to the question "what works?" to address the problem. Similar interventions have widely differing results in various settings, owing to local context, the causes and severity of malnutrition, and the capacity for program implementation.

Impact evaluations of World Bank–supported programs, which are generally large-scale, complex interventions in low-capacity settings, show equally variable results. The findings confirm that it should not be assumed that an intervention found effective in a randomized medical setting will have the same effects when implemented under field conditions. However, there are robust experimental and quasi-experimental methods for assessing impact under the difficult circumstances often found in field settings.

The relevance and impact of nutrition impact evaluations could be enhanced by collecting data on service delivery, demand-side behavioral outcomes, and implementation processes to better understand the causal chain and what part of the chain is weak. It is also important to better understand the distribution of impacts, particularly among the poor, and to better document the costs and effectiveness of interventions.

High levels of child malnutrition in developing countries contribute to mortality and present long-term consequences for the survivors. An estimated 178 million children under age five in developing countries are stunted (low height for age) and 55 million are wasted (low weight for height). Malnutrition makes children more susceptible to illness and strongly affects child mortality. Beyond the mortality risk in the short run, the developmental delays caused by undernutrition affect children's cognitive outcomes and productive potential as adults. Micronutrient deficiencies—of vitamin A, iron, zinc, and iodine, for example—are also common and have significant consequences.

Progress in reducing childhood malnutrition in developing countries has been slow. More than half of these countries are not on track to achieve the Millennium Development Goal of halving the share of children who are malnourished (low weight for age) by 2015. The food-price and financial crises are making achievement of this goal even more elusive.

The World Bank has recently taken steps to expand its support for nutrition in response to the underlying need and the increased urgency added by the crises.

What Do We Know about Reducing Malnutrition?

The increased interest and resources focused on the problem of high and potentially increasing rates of undernutrition raise a critical question: what do we know about the causes of malnutrition and the interventions most likely to reduce it?

The medical literature points to the need to intervene during gestation and the first two years of life to prevent child malnutrition and its consequences. It suggests that investments in interventions during this window of opportunity among children under two are likely to have the greatest benefits.

Recently published meta-analyses of the impact evaluation literature point to several interventions found effective for reducing undernutrition in specific settings. But there is a limit to how much these findings can be generalized, particularly in the context of large-scale government programs most likely to be supported by the World Bank. The meta-analyses tend to disproportionately draw on the findings of smaller, controlled experiments. There are few examples of evaluations of large-scale programs, over which there is less control in implementation. The meta-analyses also tend to focus on average impacts and generally do not explain the magnitude or variability of impacts across or within studies. Very few of the evaluations reviewed address the programmatic reasons why some interventions work or don't work; moreover, few assess the cost-effectiveness of interventions.

Objectives of the Review

This paper reviews recent impact evaluations of interventions and programs to improve child anthropometric outcomes—height, weight, and birthweight—with an emphasis on both the findings and the limitations of the literature and on understanding what might happen in a nonresearch setting. It further reviews the experience and lessons from evaluations of the impact of World Bank–supported programs on nutrition outcomes.

Specifically, the review addresses the following four questions:

1. What can be said about the impact of different interventions on children's anthropometric outcomes?
2. How do these findings vary across settings and within target groups, and what accounts for this variability?
3. What is the evidence of the cost-effectiveness of these interventions?
4. What have been the lessons from implementing impact evaluations of Bank-supported programs with anthropometric impacts?

Although many different dimensions of child nutrition could be explored, this report focuses on child anthropometric outcomes—weight, height, and birthweight. These are the most common nutrition outcome indicators in the literature and the ones most frequently monitored by national nutrition programs supported by the World Bank. Low weight for age (underweight) is also the indicator for one of the Millennium Development Goals.

Methodology and Scope

The Independent Evaluation Group systematically reviewed 46 nutrition impact evaluations published since 2000. These evaluations assessed the impact of diverse interventions—community nutrition programs, conditional and unconditional cash transfers, early child development programs, food aid, integrated health and nutrition services, and de-worming.

All the evaluations used research designs that compared the outcomes among those affected by the project with the counterfactual—that is, what would have happened to a similar group of people in the absence of the intervention. About half of the evaluations used randomized assignment to create treatment and control groups; the remainder used matching and various econometric techniques to construct a counterfactual.

Among the 46 evaluations, 12 assessed the impact of World Bank–supported programs on nutrition outcomes in eight countries. The broader review relies on the analysis of the published impact evaluations as the main source of data, but for these 12 evaluations, project documents and research outputs were reviewed and World Bank staff, country officials, and the evaluators and researchers who conducted the studies were interviewed.

Findings

A wide range of interventions had a positive impact on indicators related to height, weight, wasting, and low birthweight.

There were a total of 10 different outcome indicators for the four main anthropometric outcomes. A little more than half of the evaluations addressing a height-related indicator found program impacts on at least one group of children, and this was true for about the same share of interventions aimed at improving weight-related and wasting-related (low weight for height) indicators. About three-quarters of the 11 evaluations of interventions that aimed at improving birthweight indicators registered an impact in at least one specification, including five of seven micronutrient interventions.

There was no clear pattern of impacts across interventions—in every intervention group there were examples of programs that did and did not have an impact on a given indicator, and with varying magnitudes.

Evaluations of the nutritional impact of programs supported by the World Bank, which are generally large scale, complex, and implemented in low-capacity settings, show equally variable results. Even controlling for the specific outcome indicator, studies often targeted children of different age groups that might be more or less susceptible to the interventions. It is thus difficult to point to interventions that are systematically more effective than others in reducing malnutrition across diverse settings and age groups.

Differences in local context, variation in the age of the children studied, the length of exposure to the

intervention, and differing methodologies of the studies accounted for much of the variability in results.

Context includes factors such as the level and local determinants of malnutrition, differences in the characteristics of beneficiaries (including age), the availability of service infrastructure, and the implementation capacity of government. Outside a research setting, in the context of a large government program, many things can go wrong in service delivery or demand response that can compromise impact. Beyond this, social factors, such as the status of women or the presence of civil unrest, can affect outcomes.

These findings underscore the conclusion that it should not be assumed that an intervention found effective in a randomized controlled trial in a research setting will have the same effects when implemented under field conditions in a different setting. The findings also point to the need to understand the prevailing underlying causes of malnutrition in a given setting and the age groups most likely to benefit when selecting an intervention. Further, to improve performance, impact evaluations need to supplement data measuring impact with data on service delivery and demand-side behavioral outcomes to demonstrate the plausibility of the findings, to understand what part of a program works, and to address weak links in the results chain.

Evidence on the distribution of nutrition impacts—who is benefiting and who is not—and on the cost-effectiveness of interventions is scant.

Just because malnutrition is more common among the poor does not mean that children living in poverty will disproportionately benefit from an intervention, particularly if acting on new knowledge or different incentives relies on access to education or quality services. Fewer than half of the 46 evaluations measured the distribution of impacts by gender, mother's education, poverty status, or availability of complementary health services. Only nine evaluations assessed the impacts on nutritional outcomes of the poor compared with the nonpoor. Among the evaluations that did examine variation in results, several found that the children of better-educated mothers or children living in better-off communities are benefiting the most.

Bank-supported cash transfers, community nutrition, and early child development programs in six of eight countries had some impact on child anthropometric outcomes.

Of the 12 impact evaluations of Bank support, 11 were of large-scale government programs with multiple interventions and a long results chain. Three-quarters of the evaluations found a positive impact on anthropometric outcomes of children in at least one age group, although the magnitude of the impact was in some cases not large or applied to a narrow age group. Most of the impact evaluations involved assessment of completely new programs and involved World Bank researchers. Most used quasi-experimental evaluation designs, and two-thirds assessed impact after—at most—three years of program implementation. Only half of the evaluations documented the distribution of impacts, and only a third presented information on the costs of the intervention (falling short of cost-effectiveness analysis). In two of the countries (Colombia and the Philippines) the evaluations likely had an impact on government policy or programs.

Lessons

A number of lessons for development practitioners and evaluators arose from the review of impact evaluations of World Bank nutrition support.

For task managers:

- Impact evaluations of interventions that are clearly beyond the means of the government to sustain are of limited relevance. The complexity, costs, and fiscal sustainability of the intervention should figure into the decision as to whether an impact evaluation is warranted.

Photo by Julio Pantoja, courtesy of the World Bank Photo Library.

- Impact evaluations are often launched to evaluate completely new programs, but they may be equally or even more useful in improving the effectiveness of ongoing programs.
- There are methods for obtaining reliable impact evaluation results when randomized assignment of interventions is not possible for political, ethical, or practical reasons.

For evaluators:

- In light of the challenges of evaluating large-scale programs with a long results chain, it is well worth the effort to assess the risks to disruption of the impact evaluation ahead of time and identify mitigation measures.

- The design and analysis of nutrition impact evaluations need to take into account the likely sensitivity of children of different ages to the intervention.
- For the purposes of correctly gauging impact, it is important to know exactly when delivery of an intervention took place in the field (as opposed to the official start of the program).
- Evaluations need to be designed to provide evidence for timely decision making, but with sufficient elapsed time for a plausible impact to have occurred.
- The relevance of impact evaluations for policy makers would be greatly enhanced if they documented both the effects and costs of nutrition programs and interventions.

In sum, in approaching the impact evaluation literature and the conduct of nutrition impact evaluations, we should not be asking simply, "What works?" but rather "Under what conditions does it work, for whom, what part of the intervention works, and for how much?" These are important questions that managers should be asking in reviewing the literature; addressing them will also improve the relevance and impact of nutrition impact evaluations.

Chapter 1

Photo by Jamie Martin, courtesy of the World Bank Photo Library.

EVALUATION HIGHLIGHTS

- Malnutrition is widespread among children in developing countries, raising morbidity and mortality.
- Impact evaluations can provide insights about effective interventions to reduce malnutrition, though the findings are variable.
- The World Bank is ramping up its nutrition response and its impact evaluation efforts.
- This report reviews the findings of recent nutrition impact evaluations, the experience of evaluations of the nutrition impact of Bank support, and the use of the evaluation results to improve outcomes.

Introduction

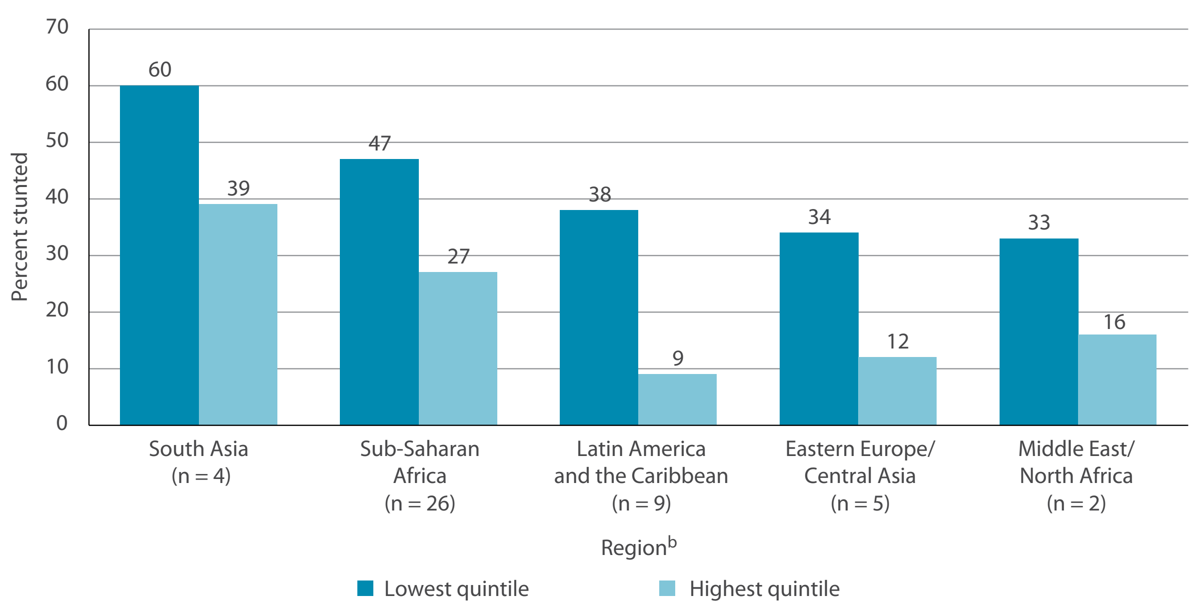
This report reviews recent impact evaluations of interventions and programs that seek to reduce child malnutrition as measured by low anthropometric outcomes. The objective is to distill lessons on effective approaches and to improve the relevance of nutrition impact evaluations of World Bank–sponsored programs.

The Heavy Toll of Malnutrition in Developing Countries

High levels of child malnutrition in developing countries contribute to high mortality and have long-term consequences for the survivors. An estimated 178 million children under the age of five in developing countries (32 percent) are stunted (low height for age), and 55 million (10 percent) are wasted (low weight for height) (Black and others 2008).[1] Within countries, *undernutrition*—in terms of stunting, wasting, and underweight—is far worse among the poor than among the nonpoor (figure 1.1). Increasing levels of underweight (low weight for age), stunting, or wasting make children more susceptible to death from common infectious diseases that do not affect better-nourished children (Caulfield and others 2006). Beyond the mortality risk, the developmental delays caused by undernutrition affect children's cognitive development and productive potential as adults. Maternal and child undernutrition are estimated to be the underlying cause of 3.5 million deaths annually (Black and others 2008, p. 243).

One-third of the children under five are stunted and one child in ten is wasted—the poor are most affected.

FIGURE 1.1 Stunting[a] among Children under Five by Developing Region and Socioeconomic Status



Source: Van de Poel and others 2008, based on the most recent Demographic and Health Survey data for 47 countries.

a. The percentage of children less than –2 standard deviations below the median height of children of the same age in the World Health Organization reference population.

b. Regional medians for South Asia, Europe and Central Asia, and the Middle East and North Africa are calculated by the Independent Evaluation Group, based on table 2 of Van de Poel and others 2008. East Asia is not presented because there was only one country (Cambodia) from that Region. The levels of undernutrition by quintile in the two North African countries (Egypt and Morocco) were remarkably similar.

Micronutrient deficiencies are also common among children in developing countries and have significant consequences (Caulfield and others 2006, p. 552–54). Vitamin A deficiency, estimated to affect from 1 percent to 40 percent of children under five, is a preventable cause of blindness and raises the severity and mortality risk of infectious diseases such as measles, diarrhea, and malaria. Iron deficiency anemia, which affects 22 percent–76 percent of children under five, can cause neurological impairment and a reduction in immune function. Zinc deficiency affects 7 percent–79 percent of children. It retards growth and increases susceptibility to infection. Iodine deficiency can lead to mental retardation and impaired physical growth, reducing the earnings of affected children when they reach adulthood.

> Malnutrition affects cognitive development and long-run productive potential and raises a child's risk of dying.

Although the overwhelming focus of public policy for child malnutrition in developing countries has been on undernutrition, childhood obesity is a growing problem and carries different health risks. Average overweight (high weight for height) among preschool children in developing countries is on the order of 3 percent, but is substantially higher in some regions and subregions.[2] The United Nations Children's Fund (UNICEF) has identified 20 countries in which more than 5 percent of preschool children are overweight, a prevalence that often exceeds the share of children who are wasted (UNICEF 2007). Childhood obesity is associated with high blood pressure, diabetes, and respiratory illness in childhood. To the extent that obese children become obese adults, they are at increased risk of chronic diseases such as diabetes, hypertension, and cardiovascular disease (De Onis and Blössner 2000).

> More than half of countries are not on track to halve the share of children who are underweight by 2015.

Slow progress in reducing undernutrition has been set back by the global food and financial crises. According to the *Global Monitoring Report 2009,* more than half of the countries with available data are not on track to achieve the Millennium Development Goal (MDG) of halving the share of children who are malnourished (underweight) by 2015 (World Bank 2009a, Annex, MDG 1, figure 4). None of the Sub-Saharan African countries with available data is on track to reduce the under-five mortality rate by two-thirds—a goal that is heavily influenced by high malnutrition (World Bank 2009a, Annex, MDG 4, figure 2). The food price and financial crises will push many more people into poverty, exacerbating malnutrition and making the MDGs even more difficult to attain. The *Global Monitoring Report 2009* estimates that 1 billion people suffer from hunger, 2 billion are undernourished and 44 million more will suffer the lasting effects of childhood malnutrition in 2008 because of these crises, with implications for health, cognitive development, and, eventually, earnings (World Bank 2009a). Achieving the MDG for malnutrition will affect the ability to achieve the goals of reducing child and maternal mortality and of boosting schooling.

> The food and financial crises have set back efforts to reduce malnutrition.

The World Bank Is Ramping Up Its Nutrition Response

Following a decade of low and declining lending for nutrition, the World Bank has taken steps to expand its support. Over the decade 1997–2006, the share of World Bank lend-ing for nutrition objectives declined, from 12 percent to 7 percent of approved projects managed by the health, nutrition, and population (HNP) sector (IEG 2009, p. 18).[3] However, *Repositioning Nutrition as Central to Development* in 2006 (World Bank 2006a) and the 2007 strategy for HNP (World Bank 2007a) renewed the commitment to reduce malnutrition and to pilot innovations in service delivery in Latin America and the Caribbean (World Bank 2009, p. 22).[4] More than 20 impact evaluations of interventions to reduce undernutrition are under way as part of the Development Impact Evaluation Initiative (DIME) coordinated by the Research Department of the World Bank (World Bank 2009c).[5]

Beyond this, in May 2008 the Bank's Board provided $1.2 billion in rapid financing through the Global Food Price Crisis Response Program, offering access under fast-track procedures to International Development Association (IDA)/International Bank for Reconstruction and Development (IBRD) grants, credits, and loans and an additional $200 million in grants for the poorest and most vulnerable countries. These emergency funds had financed projects in 30 countries as of mid-March 2009 for targeted safety nets, food-for-work programs, emergency food aid distribution, and school feeding programs, among other interventions.

What Do We Know about Reducing Malnutrition?

The increased interest and resources focused on the problem of high and potentially increasing malnutrition raises the immediate question, "What do we know about the causes of malnutrition and the interventions most likely to reduce it?" Many factors determine nutrition outcomes, and the pathway connecting public policy, private behavior, and better nutrition is complex. The medical literature points to the need to intervene in the first two years of life to prevent child malnutrition and its consequences. Recent published reviews of the literature point to promising interventions, but the generalizability of the findings of such studies is limited, particularly for national nutrition programs with multiple activities and long results chains, as implemented in field settings.

Intervening early in life is key.

The first two years of life are the window of opportunity to prevent malnutrition and its consequences. At birth, children in developing countries are remarkably similar to children in well-nourished populations in their weight and length, but growth begins to falter immediately and precipitously after birth, continuing to decline for up to three years (Shrimpton and others 2001). Children's weight, given their height, begins to decline at age three months, but it eventually recovers to levels only slightly lower than those seen in well-nourished populations. However, the mean levels of stunting of young children generally do not recover; the children grow at the same rate as the reference population, but are much shorter for their age. Gestation and the first year of life are critical periods of human brain development; it is thus not surprising that there is a correlation between low birthweight (LBW) and stunting early in life and later cognitive deficits (McGregor and others 2007; Walker and others 2007). This points to the importance of intervening early to prevent stunting and its long-run consequences. It also suggests that the potential for interventions to prevent malnutrition is greatest during pregnancy and the first 24 months of life (Bhutta and others 2008; Shrimpton and others 2001; World Bank 2006a).

> Children are particularly vulnerable to malnutrition in the first years of life.

Many causal pathways lead to nutrition outcomes.

Children and their mothers become undernourished through many causal pathways. Figure 1.2 highlights both the main pathways and the channels through which public policy can affect them. It also underscores the critical role of household and individual behavior in ensuring the success of any intervention.

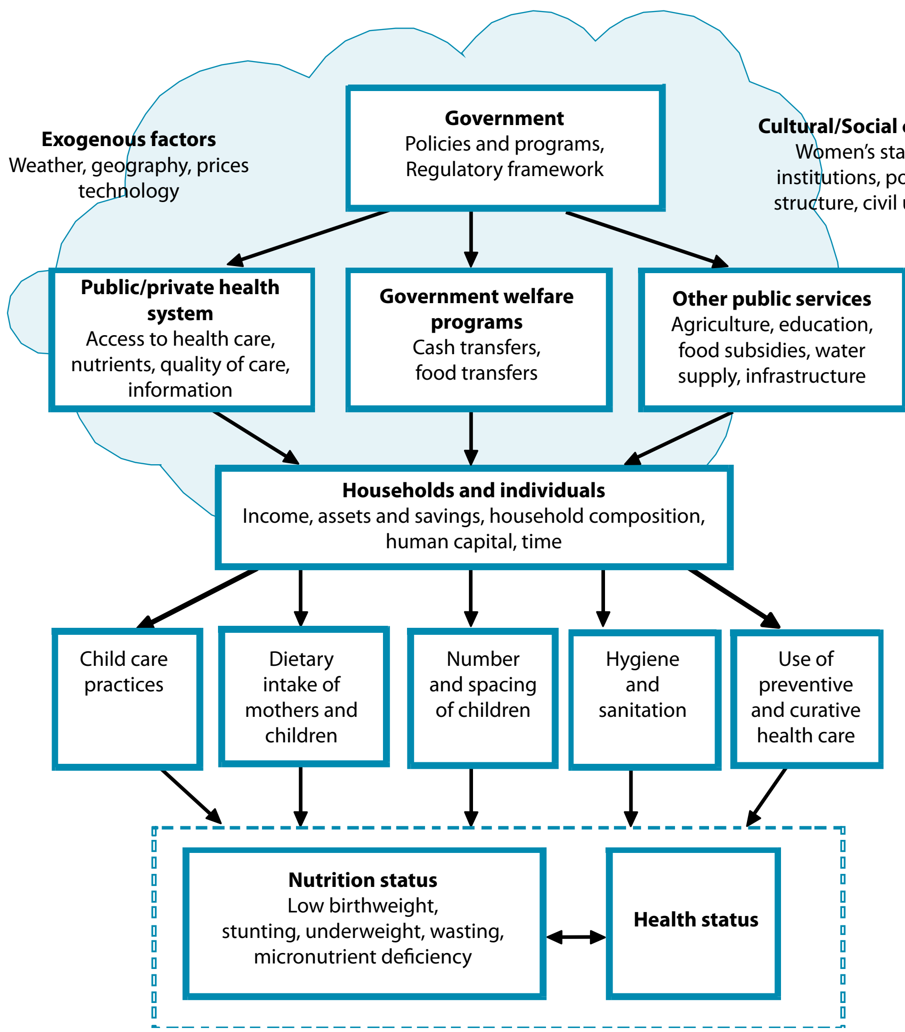
In the lower half of the figure, the immediate, proximate factors affecting child undernutrition and LBW have to do with the quality and quantity of food intake, childcare practices (such as the duration of breastfeeding and the timing of introduction of solid foods), the number and spacing of the mother's pregnancies and her own nutritional status, personal hygiene and sanitation facilities (including hygiene behaviors and water treatment), and the use of preventive and curative health care. The figure also highlights the central point that child nutritional status and health status are strongly related: low nutritional status makes children more vulnerable to illness and at higher risk of death if they become ill, and many illnesses—particularly diarrheal disease—can contribute to acute or chronic malnutrition. Further, malnutrition and infection are affected through many of the same channels.

As shown in the upper half of the figure, public policy can have an impact through government finance and regulation of many types of services—from preventive and curative health or nutrition services to safety net programs, education, agricultural information and extension, and safe water. In the background, all the actors and outcomes can be affected by exogenous factors beyond their control, such as climate (for example, drought or floods), geography, macroeconomic variables (global food or fuel prices or labor market conditions, for example), or social context (for example, the status of women, institutions, and civil unrest).

> The pathway connecting public policy to nutrition outcomes is long and complex.

These complex pathways and the numerous actors involved in implementing interventions point to a few important considerations in reviewing the literature on what works in reducing malnutrition. Because of the different local contexts in which interventions are implemented, the role of service providers and households in determining outcomes, and the lengthy results chain, the results of government nutrition programs as implemented in the field conditions of developing countries are likely to be quite different from results of randomized trials of discrete interventions in a controlled setting.

FIGURE 1.2 Pathways from Public Policy to Child Nutrition Outcomes



Sources: Authors' construction, adapted from Black and others 2008, Ruel and Hoddinott 2008, Smith and Haddad 2002, and UNICEF 1990.

First, many factors affect nutrition; we might not expect similar results across settings for a given intervention, even if it could be implemented in exactly the same way in each case. Access to nutrients can be important in some contexts, but there are populations with access to adequate food who nonetheless suffer from undernutrition because of poor feeding practices or diarrheal disease linked to poor hygiene and unsafe water. Mothers' knowledge of childcare practices may improve, but low access to health or nutrition services may prevent them from realizing the benefits of that knowledge.[6] The impact of an intervention will also depend on baseline levels of malnutrition, with a greater impact likely among those in greatest need. Thus, the measured impact of a given intervention may differ widely across settings, depending on the baseline levels of malnutrition, the root causes of the problem in that setting, and the extent to which other significant causes are working in parallel (Allen and Gillespie 2001). An intervention is also likely to have differential impacts on nutritional status of different groups of people *within* countries, depending on context.

> The impact of public policy on nutrition outcomes depends on local context . . .

Second, the effectiveness of any intervention is likely to depend on the behavior of two groups of people—service providers and households. The quality of service delivery involves incentives and decisions by health workers, be they in government, the private sector, or a nongovernmental organization (NGO). Are they trained? Will they come to work? Will their advice be good? Will they have the drugs they need, the fuel for transport, and other complementary inputs? To what extent, in effect, will the intervention be implemented as designed?

. . . and on the behavior of service providers and households.

Household and individual behaviors also affect impact. Will households participate in the program? If so, which households, and which household members? Will they change their behavior? It is rare to find a public program or intervention that does not substantially involve behavioral aspects on both the supply and demand side.[7] But in most instances the effectiveness of public programs in reducing malnutrition hinges to some extent on the ability of providers to deliver services effectively and on the extent to which the intervention enables households and individuals to make better choices. *Thus, in trying to understand whether an intervention works and why or why not, it would be important to understand whether both provider and household behaviors have changed in a way that is compatible with the intervention* (Victora, Habicht, and Bryce 2004).

Third, the causal chain between public policy and nutrition outcomes is a long one. Randomized clinical trials of specific nutrition interventions in controlled experiments—referred to in the public health literature as ***efficacy studies***—generally have a short, direct link between the intervention and the outcome (Victora, Habicht, and Bryce 2004). This type of evaluation can establish the technical efficacy of an intervention in controlled conditions. In contrast, the results chain for large-scale programs is longer and more complex, often including multiple interventions and implemented by government workers or contractors with their own incentives. *The data needs for understanding what works in a large-scale program—**effectiveness studies** in the public health literature—encompass information on the entire causal chain of intermediate outputs and outcomes.* Without this information, it is difficult to know how to interpret the differences in outcomes between program recipients or nonrecipients—whether the interventions were implemented as planned, whether households participated and their behavior changed, who benefited, and which parts of the program worked or did not work and why (Heckman and Smith 1995; Ravallion 2009a).

Recent meta-analyses provide limited guidance for what works in the context of large-scale nutrition programs.

The most recent comprehensive meta-analysis of the impact of nutrition interventions appeared in *The Lancet* in early 2008 (Bhutta and others 2008). The review included not only rigorous impact evaluations but also other types of published and unpublished program evaluations. The authors grouped their findings according to who was affected (mothers, newborn babies, and infants and young children), the intervention, and the strength of the evidence.

Understanding "what works" in large-scale nutrition programs requires information from the entire causal chain.

This follows on an earlier review of the efficacy and effectiveness of nutrition interventions in low-income Asian and Pacific countries (Allen and Gillespie 2001). These two meta-analyses found a number of consistent results, particularly with respect to micronutrient supplementation. Among the main findings from the 2008 review:


Photo by Mararu Goto, courtesy of the World Bank Photo Library.

- Promoting breastfeeding has been shown to have a large impact on child survival but little effect on stunting.
- Education about complementary feeding of children has been shown to increase height for age in populations

with sufficient food; the same result requires food supplements (with or without nutrition education) in populations with inadequate food.

- The case-fatality rate can be reduced by more than half by managing severe acute malnutrition following the World Health Organization guidelines.
- Iron folate supplements can increase hemoglobin in pregnant women, and micronutrients reduce the risk of LBW.

Despite the large number of studies reviewed, these conclusions were based on a much smaller group of evaluations of the same intervention that measured outcomes in the same way (Bhutta and others 2008, p. 421).[8] There was no attempt to compare the effectiveness of different interventions to achieve the same outcome.

Unfortunately, these meta-analyses provide limited guidance on what is likely to work in large-scale programs as implemented in the conditions of developing countries. Most of the studies reviewed by Bhutta and others (2008) consisted of smaller-scale, often randomized, pilot efficacy studies of single interventions; fewer than 3 percent of the interventions were assessed as part of effectiveness studies of large-scale programs. Allen and Gillespie (2001) admit that there were "few published examples of well designed evaluations of community-based nutrition interventions" (as opposed to those based in health facilities) and that "it is rare to find a rigorous evaluation which has demonstrated plausibly the net effects that are clearly attributable to a community-based nutrition intervention" (p. 69). Bhutta and others (2008) caution that the results of efficacy studies can overstate potential benefits of scaled-up interventions, as they "fail to include the reality of lower coverage and technical and logistical difficulties that hamper implementation in health systems" (p. 434).[9]

> Most of the research literature on nutrition impacts is based on randomized controlled trials.

The evidence of nutrition impact from large-scale programs with multiple interventions is more ambiguous. A recent review assessed the impact of conditional cash transfers (CCTs) on utilization of health care and on final nutritional outcomes, among other variables, using information from eight evaluations of seven programs in five countries, almost all of them in Latin America (Fiszbein and Schady 2009).[10] Most of the programs were implemented on a large scale, providing to the poorest households cash transfers that represented from 7 percent to 27 percent of per capita income, conditioned on use of health or nutrition services.

Both the additional income and the conditionality could have an impact on anthropometric outcomes. The authors concluded that there was evidence that CCTs raised the use of health and nutrition services and reduced disparities in the use of services by income group. However, the evidence of impact on final nutrition outcomes, such as child growth, was variable. Three of the four evaluations of programs in Mexico showed positive impacts on height or change in height, though not necessarily of great magnitude, and a fourth evaluation showed no long-run impact on height. Two evaluations showed a significant positive impact of the CCT on height for age, but in three cases there was no effect; in Brazil, the impact on weight for age was negative.

> Large-scale programs with many activities are evaluated less frequently.

Meta-analyses are heavily influenced by the results of randomized evaluations that shed little light on the implementation or programmatic factors that led to success or lack of it. The medical literature in particular tends to focus on the difference in mean health outcomes between treatment and control groups. Very little is typically learned about the performance of the intervention itself—what parts of the causal chain worked and what parts did not; this type of information, however, is important in understanding how to improve effectiveness. Fiszbein and Schady (2009) comment, for example, that it is not clear whether the variation and in many cases lack of results for CCTs—which generally are large-scale programs—reflect "differences in the data and estimation choices or underlying differences in population characteristics and program design or implementation" (p. 151). They speculate that the reason for lack of impact could have to do with "important constraints at the household level that are not addressed by CCTs as currently designed, perhaps including poor parenting practices, inadequate information, or other inputs into the production of . . . health" (p. 163).

> Randomized evaluations rarely provide information on what part of an intervention worked.

The usefulness of meta-analyses for those interested in understanding the impact of large-scale government nutrition programs of the type typically supported by the World Bank is further limited by their lack of focus on the range of results, on the distribution of impacts, and on cost-effectiveness. The emphasis in the meta-evaluation by Bhutta and others (2008) was on characterizing the average effect across studies, rather than on explaining the variation in results. The range of impact estimates is typically large, but the specific contexts and differences in the interventions underlying this variability are rarely discussed. The reviews are often organized to examine the impact of individual interventions; they rarely compare the impact of alternative interventions to achieve the same outcome. Meta-evaluations typically do not report on findings on the *distribution of impacts* across study subjects—that is, who benefits and who does not.[11] Further, very few studies present evidence on the *cost-effectiveness* of interventions, alone or comparatively.[12]

Objectives of This Study

As the World Bank moves to expand its efforts to address malnutrition—both by financing programs and by incorporating more rigorous impact evaluation—it is important to understand in greater detail what the impact evaluation research has found and how future nutrition impact evaluations can be made more relevant and useful for policy makers.

> This report addresses neglected issues in recently completed evaluations of impacts on child height and weight.

This report addresses four questions not addressed in the recent meta-evaluations of nutrition impact evaluations. First, what can be said about the impact of different interventions on children's anthropometric outcomes? Second, how do these findings vary across settings and within target groups, and what accounts for this variability? Third, what is the evidence of the cost-effectiveness of these interventions? Finally, what have been the lessons from implementing impact evaluations of Bank-supported programs with anthropometric impacts?

The report focuses on impact evaluations completed since 2000 that assess the impact of interventions on child anthropometric measures in developing countries. *Impact evaluations* are defined as those that measure an effect of an intervention by constructing a counterfactual—what would have happened to similar individuals in the absence of the intervention—and comparing outcomes under the counterfactual with the outcomes in the treatment group. They include evaluations using a variety of experimental and quasi-experimental methods. The report focuses on evaluations of the impact of programs on child anthropometric outcomes, including weight, birthweight, and height, because these are the most common nutrition outcome indicators in the literature and those most commonly monitored in national nutrition programs supported by the Bank. Underweight—low weight for age—is also the indicator for one of the MDGs. Finally, in contrast to the meta-evaluations of the literature, the report organizes the evidence so that the impacts of diverse programs can be compared with respect to a common outcome.[13]

Chapter 2 reviews the methodology and findings of 46 evaluations published since 2000 that measured the impact of various interventions on child anthropometry and LBW. In addition to reviewing the average effects found by these evaluations, it asks the following questions: How do results vary across studies, and what explains the variation? How are the impacts distributed across individuals? What do the results tell us about the effectiveness of specific program elements? How much did the interventions cost in relation to their impact? The review does not attempt to be exhaustive; its purpose is to shed light on these other questions that often are not addressed in the meta-evaluation literature, using a limited number of recent evaluations that assessed the impact of interventions on some of the most commonly researched nutrition outcomes.

> The report also reviews the results of and lessons from impact evaluations of World Bank nutrition support.

Chapter 3 reviews in depth the experience of a subset of the 46 impact evaluations—those linked to World Bank support for nutrition outcomes. The review of 12 nutrition impact evaluations of Bank support in eight countries addresses such issues as the relation between the project design and the impact evaluation, the use of the data, the use of routine administrative data, the role of local researchers, the impact of the evaluation results on the implementation of the program, and the impact of the evaluation on local capacity and public policy. The findings are based on a review of World Bank project documents, impact evaluation reports, and interviews with those involved (World Bank task managers, researchers, and country policy makers).

Chapter 4 summarizes the findings. It suggests that, going forward, we should not be asking simply what works in reducing malnutrition, but rather under what conditions it works, for whom, what part of the intervention works, and for how much. These are important questions that managers should be asking in reviewing the literature; addressing them will improve the relevance and utility of future nutrition impact evaluations.

Chapter 2

Photo by Curt Carnemark, courtesy of the World Bank Photo Library.

EVALUATION HIGHLIGHTS

- A wide range of interventions has been evaluated with respect to impact on child anthropometric outcomes.
- Many programs have shown positive impacts, yet the findings show great variability, even controlling for the intervention and the age of the child.
- Results are sensitive to local context, age group, duration of exposure, and evaluation methods.
- Few of the evaluations measure the distribution of impacts by gender, education, or poverty.
- Most of the nutrition impact evaluations lack evidence on outputs and intermediate outcomes; very few measure costs or cost-effectiveness.

Findings from Recent Nutrition Impact Evaluations

This chapter systematically reviews impact evaluations of interventions to improve child anthropometric outcomes in developing countries since 2000. It compares the average program impacts across evaluations as well as—where possible—the heterogeneity of impacts in the beneficiary population and the cost-effectiveness of interventions.

Most interventions have positive impacts on anthropometric outcomes in some settings and age groups, yet there is considerable variation in the results. The review finds evidence that this variation is partly explained by local context, the choice of the age group, the duration of exposure to the intervention, and the evaluation method. The evidence shows no clear pattern across interventions—in every intervention group there are examples of programs that did and did not have an impact on a given indicator. The review concludes that results are context specific and that it is not possible to point to certain interventions that are systematically more effective than others in reducing malnutrition across diverse settings.

Methodology

This review is based on 46 impact evaluations published since 2000 of interventions to improve child anthropometry and birthweight in developing countries. An *impact evaluation* is defined as one that attempts to construct a counterfactual as the basis for measuring changes in nutritional outcomes attributable to the program or intervention. Because there has already been a large recent meta-analysis of nutrition interventions (Bhutta and others 2008), this review focuses on a subset of the literature that measured the impact of interventions and programs on child anthropometric outcomes—indicators based on child weight, height, and birthweight. These are among the most common outcome indicators in World Bank–supported nutrition projects. The review assesses the impact on undernutrition; studies of obesity have not been included. The review is not intended to be comprehensive, but rather to identify a subset of the recent nutrition impact evaluations for closer examination of issues often not sufficiently covered in larger meta-analyses.

Selection criteria

The Independent Evaluation Group (IEG) conducted an online search of Pub Med, J-Stor, and Google Advance Scholar using relevant key words for the year 2000 through mid-2009. Other databases searched were the working papers and publications of the World Bank, the International Food Policy Research Institute, and the Integrated Management of Childhood Illness (IMCI) program. Evaluations that did not measure weight, height, or birthweight were excluded.[1] Evaluations of water supply and sanitation were explicitly excluded to keep the sample to a reasonable size and in light of other recent reviews of that literature (IEG 2008). Also excluded were evaluations that did not use experimental or quasi-experimental methods—such as randomization, propensity score matching, double-differencing, instrumental variables, or regression discontinuity methods—to construct the counterfactual. The final set of 46 evaluations includes 35 articles from peer-reviewed journals (76 percent), 7 World Bank working papers (15 percent), and 4 working papers from other institutions (9 percent).

All the evaluations tried to measure impact by comparing program outcomes with a counterfactual—what would have happened without the intervention.

Description of the sample of evaluations

A list of the 46 evaluations reviewed, by country, type of intervention, evaluation method, and anthropometric outcome indicators analyzed, is presented in table 2.1.

Geographic distribution and income level. The evaluations represent evidence from 25 developing countries. About half (52 percent) are of interventions in countries from Latin America and the Caribbean, 28 percent in African countries, and 20 percent in East and South Asian countries. There are no evaluations from the Middle East and North Africa or from Eastern Europe and Central Asia. About half of the evaluations (54 percent) took place in low-income developing countries; the remainder were conducted in middle-income countries.

The programs evaluated were in Latin America and the Caribbean, Sub-Saharan Africa, and East and South Asia.

TABLE 2.1 Interventions, Components, Countries, Evaluation Method, and Outcomes Analyzed					
Intervention/ program	Country	Components[b]	Source	Evaluation method[c]	Anthropometric outcomes analyzed[d]
Conditional cash transfers (9 evaluations)					
Bolsa Alimentação	Brazil	CT, F, G, M, NE, P, T	Morris and others 2004	IV[e]	HAZ, WAZ
Familias en Acción	Colombia	CT, F, G, M, NE, T	Attanasio and others 2005	PSM, DID	HAZ*, BW*
Oportunidades	Mexico	CT, F, G, M, NE, P, T	Leroy and others 2008	PSM, DID	Height*, weight*, HAZ*, WHZ*
Oportunidades	Mexico	CT, F, G, M, NE, P, T	Behrman and Hoddinott 2005	R, FE	Height*
Oportunidades	Mexico	CT, F ,G, M, NE, P	Barber and Gertler 2008	R, IV	BW*, LBW*
Oportunidades	Mexico	CT, F, G, M, NE, P, T	Rivera and others 2004	R	Height*
Oportunidades	Mexico	CT, F, G, M, NE, P, T	Gertler 2004	R	Height*, stunting
Atención a Crisis	Nicaragua	CT, F, G, M, NE, P, T	Macours, Schady, and Vakis 2008	R	HAZ, WAZ, BW, LBW
Red de Protección Social	Nicaragua	CT, G, M, N E, P, T	Maluccio and Flores 2005	R, DID	HAZ, stunting*, underweight*, wasting
Unconditional cash transfers (3 evaluations)					
Bono Solidario	Ecuador	CT	Leon and Younger 2007	IV	HAZ*, WAZ*
Bono de Desarrollo Humano	Ecuador	CT	Paxson and Schady, forthcoming	R	Height, HAZ
Child Support Grant	South Africa	CT	Agüero, Carter, and Woolard 2007	PSM	HAZ*
Community-based nutrition (8 evaluations)					
Bangladesh Integrated Nutrition Project	Bangladesh	F, G, M, NE, P	Hossain and others 2005	Matching	Stunting, underweight, wasting
Bangladesh Integrated Nutrition Project	Bangladesh	F, G, M, NE, P	White and Masset 2007/IEG 2005	PSM, other	HAZ*, WAZ*, WHZ*
World Vision programs	Haiti	F, G, M, NE	Ruel and others 2008	R	HAZ*, WAZ*, WHZ*, stunting*, underweight*, wasting*
SEECALINE[a]	Madagascar	F, G, M, NE, P	Galasso and Umapathi 2009	PSM, DID	HAZ*, WAZ*, stunting*, underweight*
SEECALINE[a]	Madagascar	F, G, M, NE, P, S	Galasso and Yau 2006	PSM	Underweight*
Programme de Renforcement de la Nutrition	Senegal	D, G, M, NE, P	Linnemayr and Alderman 2008	PSM, DID	WAZ*
Programme de Renforcement de la Nutrition	Senegal	D, G, M, NE, P	Alderman and others 2009	DID	Underweight*
Community Empowerment and Nutrition Project	Vietnam	D, G, F, NE	Schroeder and others 2002	R	HAZ, WAZ, WHZ, stunting, underweight, wasting
Early child development (4 evaluations)					
Proyecto Integral de Desarrollo Infantil	Bolivia	DC, F, G, M	Behrman, Cheng, and Todd 2004	PSM	Height, weight
Hogares Comunitarios	Colombia	DC, F, G, M	Attanasio and Vera-Hernandez 2004	IV	HAZ*, WAZ
Early Child Development	Philippines	F, G, M, NE, P, T	Armecin and others 2006	DID, PSM	HAZ, WHZ*, stunting, wasting*
Early Child Development	Uganda	D, G, NE	Alderman 2007	DID	WAZ*
Feeding/Food transfers (5 evaluations)					
School meals and take-home rations	Burkina Faso	F, THR	Kazianga, de Walque, and Alderman 2009	R, DID	HAZ, WAZ*, WHZ*
Food aid	Ethiopia	FFW, FD	Yamano, Alderman, and Christiaensen 2005	IV	Height*
Food aid	Ethiopia	FFW, FD	Quisumbing 2003	Other	HAZ, WHZ*
NGO feeding post *(Partage)*	Tanzania	F	Alderman, Hoogeveen, and Rossi 2006	IV	HAZ*, WAZ*
Vaso de Leche	Peru	FT	Stifel and Alderman 2006	IV	HAZ

(continued on next page)

TABLE 2.1 (continued)

Intervention/ program	Country	Components[b]	Source	Evaluation method[c]	Anthropometric outcomes analyzed[d]
Integrated health services (3 evaluations)					
Integrated Management of Childhood Illness	Brazil	NE	Santos and others 2001	R	Height, weight*, HAZ, WAZ*, WHZ*
Integrated Child Development Services	India	Various[f]	Das Gupta and others 2005	PSM	HAZ, WAZ
Integrated Management of Childhood Illness	Tanzania	Not clear[g]	Masanja and others 2005	Matching	Stunting*, underweight*, wasting
De-worming (3 evaluations)					
Primary school de-worming	Kenya	D, hygiene education	Miguel and Kremer 2004	R	HAZ*, WAZ
Pratham Delhi Preschool Health Program	India	D, M	Bobonis, Miguel, and Sharma 2006	R, DID	HAZ, WAZ*, WHZ*
ECD/De-worming	Uganda	D, DC, G, M, NE, P	Alderman and others 2006	R	Weight*
Micronutrient only (7 evaluations)					
Micronutrient	China	M (iron, folic acid, multiple)	Zeng and others 2008	R	BW*, LBW
Micronutrient	India	M (multiple containing 29 vitamins and minerals)	Gupta and others 2007	R	BW, LBW*
Micronutrient	Mexico	M (iron, multiple)	Ramakrishnan and others 2003	R	BW, LBW
Micronutrient	Nepal	M (multiple[h])	Osrin and others 2005	R	BW*, LBW*
Micronutrient	Nepal	M (folic acid, iron, zinc, multiple)	Christian and others 2003	R	BW*, LBW*
Micronutrient	Peru	M (zinc)	Iannotti and others 2008	R	Height, weight*, BW
Micronutrient	Zimbabwe	M (multimicronutrient[i])	Friis and others 2004	R	BW*, LBW
Others (4 evaluations)					
Nutrition education	Peru	NE	Waters and others 2006	Other	HAZ*, WAZ, stunting*, underweight
Nutrition education	Peru	NE	Penny and others 2005	R	HAZ*, WAZ*, WHZ, height*, weight*
Malaria	Mozambique	Sulphadoxine-pyrimethamine with insecticide-treated nets	Menéndez and others 2008	R	LBW*
Gardening	Thailand	Mixed gardening	Schipani and others 2002	Matching	HAZ, WAZ, WHZ, stunting, underweight, wasting

Source: IEG analysis.
Note: * = statistically significant positive impact.
a. SEECALINE = *Projet de Surveillance et Éducation des Écoles et des Communautés en Matière d'Alimentation et de Nutrition Élargi.*
b. CT = cash transfer; D = de-worming; DC = day care; F = feeding; FD = free food distribution; FFW = food for work; FT = food transfer; G = growth monitoring; M = micronutrients; NE = nutrition education; P = prenatal services; T = treatment of illness; THR = take-home rations.
c. DID = difference-in-difference; FE = fixed effects; IV = instrumental variable; Matching= simple comparison of program and nonprogram areas; Other = Heckman two-step maximum likelihood estimation; PSM = propensity score matching; R = randomized.
d. BW = birthweight; HAZ = height-for-age z-score; LBW = low birthweight; WAZ = weight-for-age z-score; WHZ = weight-for-height z-score.
e. The control is the group that was excluded because of "random administrative error."
f. Growth monitoring, supplementary feeding, preschool education, basic health services for young children, pregnant or lactating women.
g. Elements are not described in the evaluation; however, the IMCI strategy involves a number of complementary services at health facilities and communities (http://www.who.int/imci-mce/).
h. Vitamins A, E, D, B2, B12, and C; zinc; copper; selenium.
i. Vitamins A, β-carotene, thiamine, riboflavin, B6, B12, niacin, C, D, and E; zinc, copper, selenium.

Interventions evaluated. The interventions and programs assessed can be classified into several broad groups: large-scale CCTs; unconditional cash transfers (UCT); community-based nutrition; early child development; integrated health services; school feeding and food transfers; de-worming; micronutrients; and others.[2] The interventions consist of numerous component activities, as noted in table 2.1. Programs of the same type may include a different mix of activities, or cash or food transfers of different amounts; they may also be targeted to specific population groups.[3] It is important to note that all the evaluations of community-based nutrition programs and of de-worming were in low-income countries and all the evaluations of cash transfer programs (conditional and unconditional) were in middle-income countries, all but one of which were in Latin America and the Caribbean. All the cash transfer programs were targeted to women or mothers.

> The interventions can be classified by broad type, but even those of the same type involved different activities.

Anthropometric outcome indicators. The evaluations reported results across some 10 indicators related to height and weight (table 2.2). Some of the evaluations presented results for only 1 of these 10 indicators; others presented multiple indicators in the same dimension (for example, height, height-for-age z-score [HAZ], and stunting) or different dimensions (such as weight-for-age z-score [WAZ], HAZ, weight-for-height z-score [WHZ], or birthweight). The number of studies presenting results on each of the outcome indicators is shown in figure 2.1.

> Program impacts were measured for 10 anthropometric indicators of weight, height, and birthweight.

Although many of these indicators are related, they do not measure the same thing: a change in height or weight is a measure of absolute growth; HAZ, WHZ, and HAZ are relative to the median of another population; and stunting, underweight, and wasting measure the most malnourished segment of the distribution. It is possible to affect average height or HAZ, for example, without affecting the share of children stunted. To facilitate comparisons and avoid discrepancies based solely on the choice of indicator, the analysis compares results for all evaluations and interventions for each outcome indicator.

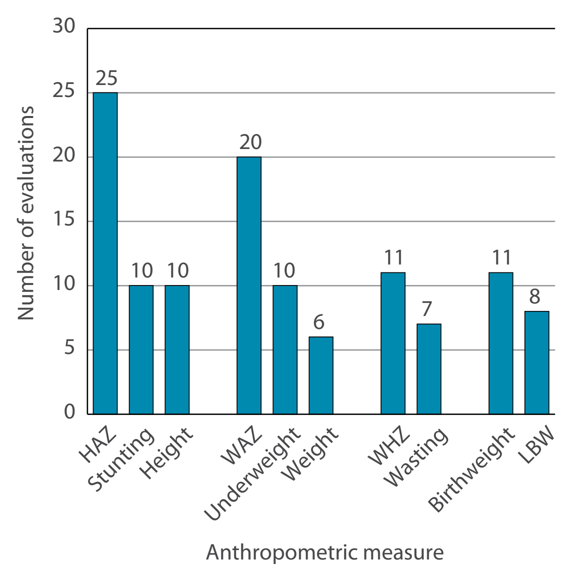
Evaluation method. Half of the 46 evaluations used an experimental design in which recipients (individuals or communities) were randomly assigned to a treatment or control group (R); the impact was measured as the difference between the outcome in the treatment and control groups.[4] The remaining evaluations used quasi-experimental methods, including propensity score matching (PSM), instrumental variables (IV), difference-in-difference (DID), or other matching to establish the counterfactual.[5] These methods are explained in appendix F.

TABLE 2.2 Definition and Interpretation of Anthropometric Indicators Used by the Nutrition Impact Evaluations

Indicator	Definition and interpretation
Height or recumbent length Weight Birthweight	These are all absolute measures of height, weight, or birthweight. Recumbent length is measured instead of height for the youngest children. Studies using these measures report the centimeters of growth in a given population, or the grams or kilograms of weight gain or birthweight. These measures are reported as mean levels in the population, with no comparison to a well-nourished reference population and no indication of the distribution of outcomes.
Height-for-age z-score Weight-for-age z-score Weight-for-height z-score	These three indicators compare a child's weight or height with the median values of a well-nourished reference population of the same age or height, and sex. The z-score measures the number of SDs above (+) or below (–) the reference population median. A child with a HAZ of –1.5 is 1.5 SDs below the median of the reference population of the same age and gender. Low HAZ is considered a measure of chronic malnutrition, while low WHZ is a measure of acute malnutrition and can change quite quickly. Low WAZ is affected by both.
Stunting Underweight Wasting	These are the percentages of children with z-score values below –2 in HAZ, WAZ, and WHZ, respectively. In other words, they are children whose measurements are more than 2 SDs below the reference population median. In the reference population, only 2.3 percent of children would normally fall below a z-score of –2. The choice of a z-score of –2 as the cutoff point is somewhat arbitrary, but these indicators are flagging the size of the group of children who are most malnourished in each dimension.
Low birthweight	Defined as the percentage of children less than 2,500 grams at birth. This is a measure of the most severely affected children.

Source: Authors, based on WHO 1995.
Note: HAZ = height-for-age z-score; SD = standard deviation; WAZ = weight-for-age z-score; WHZ = weight-for-height z-score.

FIGURE 2.1 Number of Evaluations Reporting Each of 10 Anthropometric Outcome Indicators



Source: IEG analysis.
Note: HAZ = height-for-age z-score; LBW = low birthweight; WAZ = weight-for-age z-score; WHZ = weight-for-height z-score.

Height for age and weight for height were the most commonly used indicators.

Finally, it is important to note that these impact evaluations, which primarily aim to affect anthropometric outcomes of young children, are measuring impacts over a relatively short time frame—a few years at most. The evaluations do not capture long-run impacts of undernutrition.

Half of the evaluations used an experimental design; all evaluations measured short-term nutritional impacts, not long-term consequences.

The following sections summarize and compare the impacts found in these evaluations; the extent to which they are explained by evidence of a causal chain of program inputs, outputs, and intermediate outcomes; evidence of the costs and cost-effectiveness of the interventions; and the factors underlying the variability in results.

Programmatic Impacts on Anthropometric Outcomes

The 46 impact evaluations present diverse results, in part because they assessed the impacts on groups of children of different ages and used different nutritional outcome measures. The findings below are contrasted for all interventions that present results for a common anthropometric indicator. However, the evaluations still vary in terms of the age group of the children they analyze, and this can affect the findings, in light of the specific biological windows of opportunity for affecting anthropometric outcomes.

Height, height for age, and stunting

Thirty-three evaluations were reviewed with respect to their impact on children's height, height for age, or stunting; 18 evaluations (54 percent) show positive and significant results for at least one group of children and one of these indicators; that is, either the program has significantly improved height or HAZ or reduced the proportion of stunting in program areas compared to nonprogram areas (table 2.3). However, 15 of the evaluations (46 percent) found no impact of the program on the selected height-related indicators for any of the age groups studied. Detailed findings of all evaluations of height, HAZ, and stunting are presented in appendix B.

A little more than half of the evaluations that used height indicators found a program impact.

Height/linear growth. **Most of the evidence on program impacts on height or linear growth comes from evaluations of two cash transfer programs in Latin America—one that affected height and one that did not.** Four evaluations of Mexico's CCT program, *Oportunidades,* found positive impacts on child height. In rural areas children aged 12–36 months exposed to the program were about one centimeter taller than those not exposed (Gertler 2004; Rivera and others 2004; and Behrman and Hoddinott 2005, respectively). In urban areas, children who were younger than six months at enrollment grew 1.5 centimeters more than children in the control group after two years (Leroy and others 2008). However, Ecuador's *Bono de Desarrollo Humano* (BDH), a UCT, had no impact on the height of children aged three to seven years.

A CCT program in Mexico increased height; a UCT program in Ecuador did not.

Of the five remaining programs, each a different type, only two had an impact on child height. In Ethiopia, children aged 6–24 months in the communities that received food aid grew 2 centimeters faster over 6 months, compared with the counterfactual of no aid (Yamano, Alderman, and Christiaensen 2005). In Peru, children aged 0–18 months whose mothers were exposed to nutrition education were 0.71 centimeter longer than children in the control area (Penny and others 2005). However, three programs had no impact on height—a nutrition education intervention as part of the IMCI program in Brazil (Santos and others

TABLE 2.3 Share of Evaluations with Positive Impacts on Height, HAZ, or Stunting, by Indicator and Program

Program	Height	HAZ	Stunting	Total: Height, HAZ, or stunting
Conditional cash transfers	4/4	2/5	1/2	6/8
Unconditional cash transfers	0/1	2/3	—	2/3
Community-based nutrition	—	3/4	2/4	3/5
Early child development	0/1	1/2	0/1	1/3
Feeding/food transfer	1/1	1/4	—	2/5
Integrated health services	0/1	0/2	1/1	1/3
De-worming	—	1/2	—	1/2
Micronutrient supplementation	0/1	—	—	0/1
Others	1/1	2/3	1/2	2/3
Total	**6/10**	**12/25**	**4/10**	**18/33**

Source: IEG analysis.
Note: — = There were no evaluations of the intervention with respect to this outcome variable. HAZ = height-for-age z-score. Interpretation: 4/4 = The number of evaluations that found impact (the numerator) out of the total that analyzed the outcome (the denominator).

2001), an early child development program in Bolivia, *Proyecto de Desarrollo Infantil* (PIDI) (Behrman, Cheng, and Todd 2004), and a micronutrient intervention in Peru (Iannotti and others 2008).

HAZ. **HAZ is the most frequently used indicator, analyzed in 25 studies, of which 12 reported program impacts.** As noted earlier, height for age is an indicator of chronic malnutrition. More programs can be compared in terms of their impact on HAZ than for any other indicator.

Only two of the five CCTs had an impact on HAZ, and in different age groups. Colombia's *Familias en Acción* improved HAZ of children 0–24 months old, but not of children 24–72 months (Attanasio and others 2005). In urban areas, Mexico's *Oportunidades* improved HAZ of children 0–6 months, but not of those 6–12 or 12–24 months (Leroy and others 2008). However, neither *Atención a Crisis* nor *Red de Protección Social* (RPS), both in Nicaragua, had an impact on the HAZ of children in any age group.[6] Brazil's *Bolsa Alimentação* likewise found no such impact.[7]

Conditional and unconditional cash transfer programs did not consistently affect height for age.

Two of the three UCT programs had impacts on HAZ. The South African Child Support Grants had positive impacts on HAZ on children 0–36 months,[8] as did Ecuador's *Bono Solidario* UCT program on children under five years of age, although the impact in the latter case was modest (Leon and Younger 2007). However, Ecuador's subsequent unconditional transfer program, BDH, which was better targeted to the poor, had no effect on the HAZ of children between three and seven years of age (Paxson and Schady, forthcoming).

Three of the four community nutrition programs improved HAZ. In Madagascar, the HAZ of both treatment and control groups declined, but the *Projet de Surveillance et Éducation des Écoles et des Communautés en Matière d'Alimentation et de Nutrition Élargi* (SEECALINE) program slowed the deterioration in the treatment group (Galasso and Umapathi 2009). The Bangladesh Integrated Nutrition Project (BINP) had a modest impact on HAZ of children between 6 and 23 months (IEG 2005; White and Masset 2007). In Haiti, age-based targeted interventions had a greater impact on HAZ of children in the preventive program model than on children in the traditional recuperative program (Ruel and others 2008).[9] However, the Community Empowerment Nutrition Program (CENP) in Vietnam had no impact on the HAZ of children age 5–30 months (Schroeder and others 2002).

Two of the community nutrition programs improved height for age, one showed modest results, and one had no impact.

Only one of the four feeding and food transfer (FFT) programs had an impact on HAZ. The *Partage* feeding program in Tanzania was found to have improved the HAZ of children under five (Alderman, Hoogeveen, and Rossi 2006). However, three other primarily food transfer programs did not improve HAZ: food distribution and food for work (FFW) in Ethiopia on the HAZ of children aged 0–9 years (Quisumbing 2003);[10] school meals and take-home rations (THRs) in Burkina Faso on the HAZ of children 6–60 months (Kazianga, deWalque, and Alderman 2009); and the *Vaso de Leche* program in Peru on the HAZ of children 0–59 months (Stifel and Alderman 2006).

Among the early child development programs that measured HAZ, one had a sizable impact and the other had none. Colombia's *Hogares Comunitarios* early child development program had an impact on HAZ of children six years old and younger (Attanasio and Vera-Hernandez 2004).[11] Participation in the program (captured by current attendance), the months in the program, and program exposure (months in program adjusted for age) all had positive impacts on HAZ. However, enhancements to the early child development program in the Philippines had very little impact on HAZ; it worsened in both program and nonprogram areas among children two to seven years of age (Armecin and others 2006).

> An early child development program in Colombia had a large impact on height for age, but one in the Philippines did not.

De-worming interventions had a modest impact on HAZ in one case and no impact in the other. Mass de-worming of school children 6–18 years old in Kenya, accompanied with hygiene education, produced a small and marginally significant difference in the HAZ of children in the treatment group compared with the controls (−1.13 versus −1.22, respectively) (Miguel and Kremer 2004).[12] In India, a de-worming intervention of a similar design—but including iron supplementation for the treatment group and vitamin A for the treatment and control—had no impact on HAZ of children between the ages of two and six years (Bobonis, Miguel, and Sharma 2006).

In Peru, a nutrition education intervention improved the HAZ of children 0–18 months by about 0.3 (Penny and others 2005; Waters and others 2006). However, two other programs—the Integrated Child Development Services (ICDS) program in India and mixed gardening in Thailand—had no impact on HAZ.[13]

Stunting. **Stunting is analyzed in 10 evaluations, 5 of which report program impacts.** Half of the four community nutrition programs had an impact on stunting. Madagascar's SEECALINE program reduced stunting by about 3 percent (Galasso and Umapathi 2009). The World Vision community nutrition program in Haiti reduced stunting among children in the preventive model compared with the traditional recuperative model (Ruel and others 2008). However, neither the BINP in Bangladesh (Hossain and others 2005) nor the CENP in Vietnam (Schroeder and others 2002) had an impact on stunting.

Similarly, among CCT programs, the RPS program in Nicaragua reduced stunting by 5.2 percentage points among children younger than five years of age (Maluccio and Flores 2005), but Mexico's *Oportunidades* had no impact on stunting of children 12–36 months old (Gertler 2004).

> Community nutrition programs in Madagascar and Haiti reduced stunting; those in Bangladesh and Vietnam did not.

Among the remaining programs, the nutrition education program in Peru prevented 11.1 cases of stunting per 100 children age 0–18 months, according to one evaluation (Waters and others 2006), whereas in Tanzania stunting declined more in the IMCI integrated health districts than in non-IMCI districts among children under five years of age between 1999 and 2002 (Masanja and others 2005). However, the enhanced Philippines early child development program had mixed impacts on children age two to seven years compared to children in nonprogram areas with the standard program (Armecin and others 2006)[14] and the gardening intervention in Thailand had no impact on stunting (Schipani and others 2002).

Weight, weight for age, and underweight

Twenty-eight evaluations were reviewed with respect to program impact on children's weight, weight for age, or underweight. Seventeen (61 percent) reported an impact on at least one of these indicators in children of at least one age group (table 2.4). One evaluation in Brazil found negative program impact (Morris and others 2004); the remaining 10 (36 percent) report no significant program effects on the selected weight-related indicator. Detailed findings of the evaluations reporting results on weight, WAZ, and underweight are in appendix C.

Weight. **Five of six evaluations found positive program impacts on the weight of children in different age groups in diverse programs.** The *Oportunidades* CCT program in urban Mexico improved the weight of children aged zero to six months at the time of enrollment by 0.77 kilogram; the weight of children from the lowest-income group also increased (Leroy and others 2008). The IMCI nutrition education component in Brazil raised the weight of children 12–18 months but not that of children 0–6 and 6–12 months of age (Santos and others 2006). Periodic de-worming of Ugandan preschool children aged one to seven years increased their weight by 10 percent per year when given twice a year, and by 5 percent when given annually (Alderman and others 2006). In Peru, nutrition education raised the weight of children in the intervention area by 0.199 kilogram compared with children in the control area (Penny and others 2005), and a micronutrient-supplementation program raised the weight of children under 12 months by 0.58 kilogram (Iannnotti and others 2008). However, the PIDI early child development program in Bolivia had no impact on children's weight in any age group (6–24, 25–36, 37–41, 42–58, and >59 months), even though the intervention included feeding (Behrman, Cheng, and Todd 2004).

TABLE 2.4 Share of Evaluations with Positive Impacts on WAZ, Underweight, or Weight, by Indicator and Program

Program	Weight	WAZ	Underweight	Total: Weight, WAZ, or underweight
Conditional cash transfers	1/1	1/3[a]	1/1	2/4[a]
Unconditional cash transfers	—	0/1	—	0/1
Community-based nutrition	—	4/5	4/6	6/8
Early child development	0/1	1/2	—	1/3
Feeding/food transfer	—	2/2	—	2/2
Integrated health services	1/1	1/ 2	1/1	2/3
De-worming	1/1	1/ 2	—	2/3
Micronutrient supplementation	1/1	—	—	1/1
Others	1/1	1/3	0/2	1/3
Total	**5/6**	**11/20**	**6/10**	**17/28**

Source: IEG analysis.
Note: — = There were no evaluations of the intervention with respect to this outcome variable. Interpretation: 1/1 = The number of evaluations that found impact (the numerator) out of the total that analyzed the outcome (the denominator).
a. In addition to these positive results, an additional evaluation (in the denominator) found a negative impact of Brazil's *Bolsa Alimentação* on WAZ of children seven years of age or younger (Morris and others 2004).

WAZ. **Evidence of program impact on WAZ comes from almost all intervention types, but the largest group represented is community-based nutrition programs**. Four of the five programs improved WAZ: the BINP in Bangladesh, SEECALINE in Madagascar, the World Vision nutrition program in Haiti, and the *Programme de Renforcement de la Nutrition* (PRN) in Senegal. However, the CENP community-based nutrition program in Vietnam had no impact on WAZ. BINP had a modest impact (0.07 to 0.09) on WAZ of children aged 6–23 months (IEG 2005; White and Masset 2007). SEECALINE increased the WAZ of children under five years by 0.15 to 0.22 (Galasso and Umapathi 2009). The Haiti program raised the WAZ of children 12–41 months in preventive communities by 0.24, compared with children in the recuperative communities (Ruel and others 2008). Senegal's PRN increased WAZ for children 0–6 months, but not for children aged 0–36 months (Linnemayr and Alderman 2008).

Four of five community nutrition programs improved weight-for-age scores.

Both of the food transfer programs that measured WAZ had an impact. In Burkina Faso, take-home rations (THRs) at primary school improved the WAZ of preschool children in school-age children's homes, but neither THR nor school feeding improved WAZ for school-age children (box 2.1). In Tanzania, presence of a *Partage* feeding post in the community was associated with higher WAZ (Alderman, Hoogeveen, and Rossi 2006).

In contrast, two of four cash transfer programs had no impact on WAZ, and in one CCT, WAZ actually worsened. Nicaragua's *Red de Protección Social* CCT improved the WAZ of children under six years of age (Maluccio and Flores 2005). However, *Atención a Crisis*, another Nicaraguan CCT, had no impact on the WAZ of children of any age group (Macours, Schady, and Vakis 2008), nor did Ecuador's UCT, *Bono Solidario* (Leon and Younger 2007). However, each additional month of exposure to Brazil's *Bolsa Alimentação* CCT was associated with a 0.13 lower WAZ than that observed in children of the same age in the control group (Morris and others 2004).[15]

Similarly, the impact of two early child development programs on WAZ varied. The program in Uganda raised the WAZ of children less than one year of age; no program effect was found in WAZ of children 12–24 months, 24–36 months, 36–48 months, or >48 months, however (Alderman 2007). The author noted that one would expect the younger children to experience the greatest impact because their mothers were exposed to the intervention during pregnancy. However, the *Hogares Comunitarios* early child development program in Colombia had no impact on WAZ of children 0–72 months, even though food was distributed as a component (Attanasio and Vera-Hernández 2004).

De-worming of preschool children in India improved WAZ, but de-worming of school-age children in Kenya did not. In India, the de-worming program brought about a 0.31 improvement in WAZ for children between two and six years of age, which is equivalent to an average weight gain of 0.5 kilogram (Bobonis, Miguel, and Sharma 2006).

BOX 2.1 The Impact of School-Based Feeding Interventions in Burkina Faso on School-Age and Preschool Children

A school feeding program implemented in Burkina Faso offered two interventions: school meals and take-home rations (THR). The school meals component was a lunch provided daily to attending students. The THR component was a 10-kilogram bag of cereal flour to girls, given every month conditional on a 90 percent attendance rate. The program targeted school-age children and therefore the recipients of the school meal and THR were children aged 6–15 years.

Kazianga, de Walque, and Alderman (2009) evaluated several schooling and health outcomes of these school-age children as well as the impact of this program on the nutritional status of preschool children in the same households. The underlying assumption is that the dry THRs issued to school-age children would increase food availability and hence improve the nutritional status of preschool children in the same household. The assumption in the school meals case is that the preschool children at home would receive more food than would have been the case had their older siblings not participated in the school meals program.

The evaluation found that in the THR villages, WAZ increased by 0.36 for preschool children, but there was no impact on school-age children. In the school meals villages, there was an impact on WAZ of school-age children, but not on preschool children. There was no impact on HAZ of either group of children in either program, although WHZ increased for preschool children in the THR villages.

Source: Kazianga, de Walque, and Alderman 2009.

However, there was no impact of de-worming on WAZ of school children aged 6–18 years in Kenya (Miguel and Kremer 2004).[16]

> The impact of CCTs, early child development programs, and de-worming on weight for age was variable.

Of the remaining four programs, only two had an impact in raising WAZ, and one of those is in question. The nutrition education component of the Brazil IMCI program improved WAZ among children 12–18 months, but not among children 0–6 or 6–12 months (Santos and others 2001); the ICDS health intervention in India found no impact on the WAZ of preschool children in the mid-1990s (Das Gupta and others 2005). In Peru, a nutrition education program roughly halved the (negative) WAZ of children age 18 months in the intervention area compared with children in the control area (mean values of –0.34 and –0.62, respectively) (Penny and others 2005). However, using the same data set, a second evaluation found that this impact disappears when other maternal and household characteristics are controlled for in a multivariate regression analysis (Waters and others 2006).

Underweight. **Six of the ten studies that investigated underweight are community-based nutrition programs and three of the six programs had an impact.** At the individual child level, Madagascar's community-based SEECALINE nutrition program reduced underweight among children younger than five years of age by 5.2–7.6 percentage points (Galasso and Umapathi 2009). At the community level, an additional year of exposure reduced underweight rates for children aged 0–6 months by about 8 percent and for children 7–12 months by 4 percent; two additional years of exposure reduces underweight by 8 percent in all age groups (Galasso and Yau 2006). However, the impacts varied according to the child's age when the intervention started: reductions in underweight for children aged 12–36 months are observed only after two extra years of exposure. Senegal's PRN community nutrition program (Alderman and others 2009) and the World Vision nutrition program in Haiti (Ruel and others 2008) both reduced underweight among younger children in program villages, compared to children in nonprogram villages. However, neither the Vietnam CENP (Schroeder and others 2002) nor the Bangladesh BINP community nutrition program had an impact on underweight (Hossain and others 2005).[17]

Among the remaining programs evaluated on underweight, two had an impact and two did not. In Nicaragua, the RPS CCT program reduced underweight of children 0–60 months to 9.8 percent in the program areas, and underweight increased to 16.6 percent in nonprogram areas (Maluccio and Flores 2005).[18] The Tanzania IMCI program also reduced underweight in program areas (Masanja and others 2005). However, neither mixed gardening in Thailand (Schipani and others 2002) nor nutrition education in Peru (Waters and others 2006) was found to have had an impact on underweight.

Weight for height and wasting

Weight for height and wasting are not as commonly measured as other anthropometric indicators. Only 14 of the 46 evaluations (30 percent) selected for this review presented impacts on WHZ or wasting (table 2.5). The detailed findings of these studies are in appendix D. Wasting is less prevalent than stunting and underweight.[19] In addition, a child's WHZ can change in a very short time because of acute illness, for example, which can easily overwhelm program effects. The community-based nutrition evaluations were most likely to measure WHZ or wasting (half of them did so), but only two of the nine evaluations of CCTs reported results on one of the two outcomes. Surprisingly, only two of the food-based programs measured WHZ, and none measured wasting, even though this type of intervention conceivably could have important short-run impacts on weight.

Weight for height and wasting are not often measured in the impact evaluation literature.

WHZ. **Only one of the three community-based nutrition programs that measured WHZ had an impact on it.** The World Vision community nutrition programs in Haiti—with relatively high levels of wasting—raised the WHZ of children in the preventive communities by 0.24 compared with the children in the recuperative communities (Ruel and others 2008). However, the community-based programs in Bangladesh (BINP) (IEG 2005; White and Masset 2007) and Vietnam (CENP) (Schroeder and others 2002) had little or no impact.

Both of the food aid programs that measured WHZ had an impact on it. In Ethiopia, food distribution raised the WHZ of children zero to five and five to nine years of age in high-asset households, and FFW had a similar impact on young children in low-asset households (Quisumbing 2003). The THR program in Burkina Faso raised WHZ of children 12–60 months by 0.33 after about a year (Kazianga, de Walque, and Alderman 2009). However, the result is significant only at the 10 percent level, and it disappears when the sample includes all children from 6–60 months. The school meals component of the program had no impact on WHZ.

Only one evaluation each measured WHZ for a CCT program, an early child development program, integrated health services, or de-worming interventions, but all four of these programs had an impact on WHZ. In urban areas, Mexico's CCT, *Oportunidades,* raised WHZ by 0.47 among children 0–6 months old in program areas, but not for those aged 6–12 or 12–24 months (Leroy and others 2008). The enhanced early child development program in the Philippines had predominantly positive impacts on the WHZ of children of different ages (2, 3, 4, 5, and 6 years) for different durations of exposure in the program (4–12, 13–16, and >17 months) (Armecin and others 2006).[20] A de-worming intervention in India on children between the ages of 24 and 72 months raised the WHZ of children by 0.52 five months after the intervention began (Bobonis, Miguel, and Sharma 2006). The nutrition education component of the IMCI program in Brazil improved the WHZ of children 12–18 months, but not those of children 0–6 and 6–12 months (Santos and others 2001). However, two other programs—nutrition education in Peru (Penny and others 2005) and mixed gardening in Thailand (Schipani and others 2002)—had no impact on WHZ. The first of these was aimed at children 0–18 months of age and the second at children between the ages of 1 and 7 years.

Wasting. **Only seven studies analyzed wasting, and only two reported program impacts.** Three of the seven were community-based nutrition programs. As was the case for WHZ, only the World Vision community-based program in Haiti, where 9 percent of children are wasted, had an impact on wasting (Ruel and others 2008).[21] Neither the CENP community-based nutrition program in Vietnam (Schroeder and others 2002) nor the Bangladesh BINP (Hossain and others 2005)[22] had an impact on wasting.

The other program that had an impact on wasting—the Philippines comprehensive early child development program—had predominantly positive program impacts on the wasting of children aged 2, 3, 4, 5, and 6 years for different durations of exposure to the program (4–12, 13–16, and >17 months) (Armecin and others 2006).[23] However, the Nicaragua CCT, RPS (Maluccio and Flores 2005), the Tanzania IMCI health program (Masanja and others 2006), and the mixed-gardening program in Thailand (Schipani and others 2002) had no impact on wasting. The finding in Nicaragua is

TABLE 2.5 Share of Evaluations with Positive Impacts on WHZ or Wasting, by Indicator and Program

Program	WHZ	Wasting	Total: WHZ or wasting
Conditional cash transfers	1/1	0/1	1/2
Community-based nutrition	1/3	1/3	1/4
Early child development	1/1	1/1	1/1
Feeding/ food transfer	2/2	—	2/2
Integrated health services	1/1	0/1	1/2
De-worming	1/1	—	1/1
Others	0/2	0/1	0/2
Total	**7/11**	**2/7**	**7/14**

Source: IEG analysis.
Note: — = There were no evaluations of the intervention with this outcome variable. Interpretation: 1/1 = The number of evaluations that found impact (the numerator) out of the total that analyzed the outcome (the denominator).

not altogether unexpected, as only 1 percent of children were wasted (less than the 2.3 percent in the reference population). The predominance of impact evaluations from Latin America, where wasting is low, may explain in part why so few of the 46 evaluations measured this indicator.

Birthweight and LBW

Micronutrient interventions dominate the programs for which birthweight impacts were measured (table 2.6). This review identified 11 recent impact evaluations of birthweight or LBW from nine countries—China, Colombia, India, Mexico, Mozambique, Nepal, Nicaragua, Peru, and Zimbabwe. Birthweight and the incidence of LBW respond to activities targeted to pregnant women, including micronutrient and energy supplements and other prenatal services aimed at improving dietary practices and living conditions (Allen and Gillespie 2001; Bhutta and others 2008). Seven of the 11 evaluations of birthweight and LBW measured the impact of micronutrient interventions; the only other interventions represented are CCTs and a single program targeting malaria. Notably, 10 of the 11 studies of birthweight or LBW had experimental (randomized) designs.[24] The detailed findings of evaluations that measured the impact on birthweight and LBW are in appendix E.

Most of the programs affecting birthweight involved micronutrient interventions, and most worked.

Five of the seven micronutrient programs had impacts on birthweight or LBW. Although the specific micronutrients provided varied across the programs, most offered multiple micronutrient supplementations during pregnancy to the treatment groups, compared with the standard folic acid and/or iron supplementations in the controls. Interventions offering multiple micronutrients in India (Gupta and others 2007) and Nepal (Christian and others 2003; Osrin and others 2005) both raised birthweight and reduced LBW. Programs in China (Zeng and others 2008) and Zimbabwe (Friis and others 2004) raised average birthweight but had no impact on LBW. However, neither a Peruvian program that offered only zinc (Iannotti and others 2008) nor a Mexican intervention that provided iron and a multiple micronutrient (Ramakrishnan and others 2003) had an impact on birthweight. It is interesting to note that the two programs with no impact on birthweight were in middle-income Latin American countries, whereas those that did were in low-income countries.

Three CCT programs measured impacts on birthweight, as did one malaria program. In the case of Mexico's *Oportunidades*, "beneficiary status predicts 127.3 g[rams] higher birth weight . . . and a 4.6 percentage point reduction in low birth weight" (Barber and Gertler 2008, p. 1409). The impacts were greater among women who spent more time in the CCT program and those who received more cash.[25] Colombia's *Familias en Acción* CCT also had an impact on raising birthweight. However, Nicaragua's *Atención a Crisis* had no impact on birthweight (Macours, Schady, and Vakis 2008).[26]

Finally, a program in Mozambique that provided two doses of sulphadoxine-pyrimethamine and insecticide-treated bednets reduced LBW among women who had had four or more pregnancies (Menendez and others 2008).[27]

Heterogeneity in Impacts

Aside from analyzing the average impacts of interventions across age groups, fewer than half of the studies examined the distribution of effects on the nutritional outcomes of different beneficiary groups—the impact on the poor and the nonpoor, the children of educated and uneducated mothers, or boys and girls. Only 40 percent (19 of the 46 evaluations) examined the variation (heterogeneity) of the impact of the interventions by characteristics other than age group. These included income and poverty or any other measure of socioeconomic status (9 evaluations), maternal education (6), gender (6), place of residence or region (3), and other characteristics (8).

TABLE 2.6 Share of Evaluations That Found Impacts on Measures of Birthweight

Program	Birthweight	Low birthweight	Total—birthweight or low birthweight
Conditional cash transfers	2/3	1/1	2/3
Micronutrient supplementation	5/7	3/6	5/7
Others—malaria	—	1/1	1/1
Total	**7/10**	**5/8**	**8/11**

Source: IEG analysis.
Note: — = There were no evaluations of this intervention for this outcome measure. Interpretation: 2/3 = Of the three evaluations that measured BW, two reported statistically significant impacts. There were no evaluations of the impact of UCT, community-based nutrition, early child development, food transfers, integrated health services, or de-worming on birthweight.

Fewer than half of the evaluations looked at the distribution of impacts.

Among the nine evaluations that examined impacts by socioeconomic status, most found that children from the poorest households benefit more than those from less poor households. Although programs often target the poorest group of the society, the relative differences in income or socioeconomic status within the targeted group affect the magnitude and significance of impacts.

Mexico's *Oportunidades* CCT program had a positive impact on height among rural children from the poorest households, but not on children from relatively better-off households (Rivera and others 2004). In urban areas *Oportunidades* also had a stronger impact on child growth (measured by both height and weight) for children from the poorest households (Leroy and others 2008). Among Ethiopian children younger than 5, food for work improved WHZ in low- but not high-asset households (Quisumbing 2003).

In contrast, free distribution of food raised WHZ of children younger than 5 in high-asset Ethiopian households, but not in low-asset households (Quisumbing 2003). Madagascar's SEECALINE, though targeted to the poorest areas, tended to benefit the nutritional status of children in better-off communities (Galasso and Umapathi 2009).[28]

Four programs had no differential impact on children's nutritional status across income groups or household wealth: Nicaragua's *Bono de Desarrollo Humano*, a UCT (Paxson and Schady, forthcoming); Uganda's early child development program (Alderman 2007); and the community nutrition programs in Bangladesh (IEG 2005) and Senegal (Linnemayr and Alderman 2008).

In Mexico and Colombia, the poorest children benefited the most.

Evaluations in Mexico and Madagascar suggest that children with more educated mothers benefit more than those with less educated mothers. The impact of Mexico's *Oportunidades* CCT on height was larger for children whose mothers had better education (Behrman and Hoddinott 2009). Madagascar's SEECALINE community-based nutrition program improved the HAZ, WAZ, and underweight of children whose mothers had secondary or higher education; the program also raised WAZ for children whose mothers had primary schooling but had no impact on children whose mothers had no education (Galasso and Umapathi 2009).

In contrast, in Colombia and India the children of the least educated mothers benefitted the most. In Colombia, the *Hogares Comunitarios* early child development program had a greater impact on the HAZ of children whose mothers had no education (Attanasio and Vera-Hernandez 2004). A de-worming program in India had a larger impact on the WHZ of children whose mothers had less than three years of schooling (Bobonis, Miguel, and Sharma 2006). Neither the Ugandan early child development program nor the Bangladesh community nutrition program (BINP) had differential program impacts on WAZ by mother's education (Alderman 2007; IEG 2005).

Children whose mothers had more education were more likely to benefit in Mexico and Madagascar, but less likely to benefit in Colombia or India.

The six evaluations that examined the differing impacts of programs by gender produced quite variable results, depending on the country and the intervention. The BDH unconditional cash transfer program in Ecuador benefited girls more than boys for several health and educational outcomes, although there were no impacts on the height of girls or boys (Paxson and Schady, forthcoming). Food for work in Ethiopia—where boys under nine have lower nutritional status than girls—appears to improve boys' WHZ more than girls', among children under five, and it improves boys' HAZ more than girls' in children between the ages of five and eight in low-asset households (Quisumbing 2003).[29] However, the gender effects depend on the modality of food aid (FFW versus free distribution of food), the age groups, household assets, and the specification; in most cases there are no gender effects of food aid. The ICDS program in India tended to improve the HAZ of boys more than girls in 1992, but there were no differences in impact by gender in 1998, nor were there any differences in impact by gender of WAZ in either year (Das Gupta and others 2005). The Indian de-worming program improved the WHZ of both boys and girls, but the magnitude of the impacts was larger and stronger for girls (Bobonis, Miguel, and Sharma 2006).

In contrast, there were no differential impacts on HAZ, WAZ, or WHZ by gender of the *Red de Protección Social* in Nicaragua (Maluccio and Flores 2005). A micronutrient program in Peru reported different impacts by gender but did not explain them (Iannotti and others 2008).

Evaluations have also looked at impacts by other beneficiary and program characteristics, such as place of residence, community infrastructure, number of prior pregnancies, anemia, or human immunodeficiency virus (HIV) status. The ICDS program tended to improve the WAZ of children from the northern (poor) region of India in 1998, but there were no differences in impact by region in 1992, nor were there any regional differences in impacts on HAZ in either year (Das Gupta and others 2005).

WHZ improved both in children who were anemic at baseline and in those who were not; however, the impact of the Pratham Delhi Preschool Program was greater for children who were anemic at baseline (Bobonis, Miguel, and Sharma 2006). The SEECALINE community-based nutrition program in Madagascar had greater impacts in villages with better proximity to a road, a hospital, electricity, and access to safe water source (Galasso and Umapathi 2009). However, Mexico's *Oportunidades*, a CCT, had no differential program impact on height by access to community infrastructure (Behrman and Hoddinott 2005).

Colombia's *Familias en Acción,* a CCT, had impact on birthweight in urban but not in rural areas (Attanasio and others 2005). A malaria intervention in Mozambique reduced incidence of LBW for women with four or more prior pregnancies (Menendez and others 2008). However, no differential impact was found by HIV status of women. Similarly, in Tanzania, there was no difference in the impact of multi-micronutrient supplementation on birthweight by HIV status of the woman (Friis and others 2004).

Understanding the Causal Chain

Impact evaluations have as an objective to be able to attribute an outcome to an intervention. If the control and treatment groups are identical in their composition and there is no attrition or crossover between groups, then any difference between outcomes in the two groups can be attributed to the program.

However, there are a number of reasons why it is not only prudent but highly advisable to document the causal chain of the program or intervention—from the inputs to outputs and intermediate outcomes. First, in the real world it is often difficult to prevent attrition, crossover, or other exogenous events (such as an economic or a political crisis) that can compromise an experimental design and confound the findings. Documenting implementation of the intervention and intermediate outputs and outcomes lends plausibility to the findings. It establishes whether the intervention was fully implemented, providing insight as to whether the impact might have been even larger had it been implemented correctly.

> Documenting the causal chain helps explain why outcomes were or were not achieved.

Second, documenting the causal chain helps explain why the anticipated outcomes were or were not achieved, the extent to which each part of the intervention was actually implemented, which part contributed the most or least to outcomes, and how impact might be increased. Lack of impact or small impacts can be the result of shortcomings in implementation, which cannot be assessed without information from the causal chain. Many nutrition interventions involve multiple activities, and managers want to understand which of these activities contributed to outcomes.

For community-based nutrition programs, for example, managers want to understand the contribution of feeding (the most expensive component) to better outcomes. In the case of CCTs, policy makers want to understand whether it was the cash transfer or the conditionality that was responsible for outcomes. There was an enormous increase in the uptake of iron supplement (ferrous sulfate) as a result of the RPS conditional cash transfer in Nicaragua in the treatment areas relative to the control areas between 2000 and 2002 (Maluccio and Flores 2005). Both stunting and underweight declined in the treatment areas relative to the controls. Despite this, there were no significant reductions in anemia between the treatment and control children over time. Rich data on the causal chain could offer an explanation for unexpected results, such as the worsening of WAZ in Brazil's *Bolsa Alimentação* program (Morris and others 2004). Greater attention to tracking intermediate outcomes and a process evaluation to assess implementation difficulties would have shed light on the causes of these counterintuitive results.

> Only about half of the evaluations documented at least one intermediate outcome.

Despite these benefits, only about half of the 46 impact evaluations (24) documented at least one intermediate outcome. The most commonly measured intermediate outcomes were micronutrient intake or status (13); illness (12); use of health care (9); dietary intake (7); and breastfeeding knowledge and practice (7).[30]

A few evaluations were able to infer the effectiveness of the different parts of the intervention by pointing to intermediate outcome indicators in the causal chain. In Senegal, the positive impact of PRN, a community-based nutrition program, on the WAZ of the youngest group of children was validated and explained by a concomitant increase in breastfeeding and weaning practices in program areas for the youngest children (Linnemayr and Alderman 2009).

Bangladesh's BINP community-based nutrition program had a small impact on nutritional outcomes, at best. Data on intermediate outcomes showed that women in the BINP areas had greater knowledge than women in control areas as a result of the program; however, for some reason they had not been able to translate that information into changes in practice that would improve nutrition outcomes (Hossain and others 2005; White and Masset 2007).

In Peru, the improvements in children's nutritional status could be explained in part by an increase in health care use in areas covered by the nutrition education program (Waters and others 2006). Colombia's *Familias en Acción* CCT program had an impact on intermediate outcomes, such as improved probability of compliance with preventive health care, lower morbidity, and improved food intakes. HAZ improved among children younger than 2 years old, but not for older children (24–48 months and >48 months), even though the food intake of the older children was improved by the program (Attanasio and others 2005). Similarly, *Atención a Crisis* in Nicaragua had an impact on dietary intakes and health care utilization, although this apparently did not lead to an impact on any of the child anthropometric indicators (Macours, Schady, and Vakis 2008).

In Bangladesh, women participating in the BINP community nutrition program acquired knowledge, but this did not change their behavior.

The Kenya primary school de-worming program included both de-worming and preventive health education, either or both of which could have accounted for the improvement in HAZ. However, because the evaluators were able to document no difference between the control and treatment groups in hygiene behavior, they argue that the nutritional outcome was likely a result of the de-worming drugs (Miguel and Kremer 2004).

Program Costs and Cost-Effectiveness

Impact evaluations provide an opportunity to measure the impact as well as the costs of programs, providing insights into both efficiency and sustainability. Cost-effectiveness analysis of specific elements of complex interventions is often constrained, however, by the fact that evaluations do not isolate the component that matters for the measured impact.

Among the 46 evaluations reviewed, only a handful documented the costs or cost-effectiveness of the interventions evaluated. In Uganda, a de-worming intervention was implemented with preschool children as part of "child health days" in the early child development program, which also offered polio inoculations and vitamin A supplementation (Alderman and others 2006). The cost of the health day event was estimated at $1.33 per child and the de-worming intervention at $0.25 per child per event.

In Kenya, a de-worming program helped avert 649 disability-adjusted life years (DALYs), equivalent to a cost of $5 per DALY averted,[31] but this value underestimated the health spillover benefits (Miguel and Kremer 2004). In Peru, after 18 months of follow-up of 338 children from birth, the nutrition education program was found to have averted 11.1 cases of stunting per 100 children in the 0- to 18-month age range. The estimated marginal cost, including external costs, training, health education materials, and extra travel and equipment, was $6.12 per child, or $55.16 per case of stunting averted (Waters and others 2006).

Three evaluations assessed the costs and benefits of the interventions by examining payoffs in the long run. The anthropometric improvements attributable to Mexico's *Oportunidades* CCT in rural areas were estimated to be equivalent to a 2.9 percent increase in lifetime earnings (Behrman and Hoddinott 2005). The present value of the investment in human capital resulting from the South Africa Child Support Grants exceeded by more than 60 percent the cost of the program (Agüero, Carter, and Woolard 2007). The benefit-cost ratio of the PIDI preschool program in Bolivia was calculated by estimating the benefits and costs to the child, assuming that he or she attained intermediate and secondary education (Behrman, Cheng, and Todd 2004). In a hypothetical setting,[32] the benefit-cost ratio is estimated to be 1.37–2.48 at a 5 percent discount rate; however, improved anthropometric outcomes were not among the benefits.


Photo by Julio Pantoja, courtesy of the World Bank Photo Library.

Accounting for the Variability in Results

When comparing results of evaluations with similar interventions on identical outcomes, the analysis of these 46 evaluations leads to the conclusion that there is enormous variability. This review finds evidence that some of the variation can be explained by differences in context, the age group studied, the duration of the intervention, and the evaluation method.

Context mediates the impact of nutrition interventions.

Impact evaluations of similar programs offer different results because of differences in context. The variability of the impacts of similar programs implemented in different countries or the same country in different periods or settings is evident for all types of interventions and anthropometric indicators. The programs have important differences that arise from baseline beneficiary characteristics, country, and program area, all of which can affect outcomes.

> The variation in nutrition impacts of the same programs can be explained by different contexts, exposure, age groups, and evaluation methodologies.

Baseline characteristics or initial conditions can affect the magnitude of the impact. The evaluation of de-worming in Uganda, for example, took place in the region with the highest burden (Alderman and others 2006); both the results and cost-effectiveness would likely be different in other parts of Uganda where the burden is less severe. A community-based nutrition program had an impact on WHZ and wasting in Haiti, with high baseline levels of both (Ruel and others 2008).

In contrast, there was no impact of nutrition education on WHZ in Peru, which could be attributable to the intervention or to the fact that it is at such a low level (less than in the reference population) (Penny and others 2005). The impact of a de-worming program in India on WHZ was higher among children with the most severe anemia at baseline (Bobonis, Miguel, and Sharma 2006). If certain interventions predominantly have an impact among children with educated mothers (as was found in several evaluations), then baseline levels of maternal education will affect the average impact.

> The impact of de-worming in India was greatest among children with the most severe anemia.

The availability of complementary infrastructure—not often measured in these evaluations—can also affect program impact. This review found systematic differences in the distribution of interventions by region. It is perhaps no accident that all the CCTs, in which transfers to the poorest people are conditioned on the use of health or education services, were in middle-income countries, where access to basic health services is not generally constraining. Even the UCT program in Ecuador, *Bono de Desarrollo Humano*, raised utilization of health care. However, in low-income countries health care is less accessible. Community infrastructure not only augments the impact of Madagascar's SEECALINE community-based nutrition program but also complements mother's education (Galasso and Umapathi 2009).[33]

Implementation capacity is another dimension of context, though the evaluations reviewed here had very little information to document the extent of implementation. Poorly implemented interventions can be indistinguishable from no intervention at all. The causal chain was rarely documented in these evaluations, but it is reasonable to expect that in some cases the lack of impact could be caused by poor implementation. The PIDI early child development program in Bolivia, for example, showed no impact on height or weight, even though the intervention provided food to the children; however, no information was available on the extent to which the food was delivered, the quality of home care and stimulation provided the children, the number of children per caretaker, or other indicators to understand to what extent the intervention was implemented as planned (Behrman, Cheng, and Todd 2004).

> Lack of impact of large-scale nutrition programs can be due to shortfalls in program implementation.

Finally, women's status can strongly condition the outcomes of nutrition programs. Most of the impact evaluations were of interventions targeted to women, on the assumption that they are the main decision makers concerning children's welfare. However, this may not always be the case. Evaluations of Bangladesh's BINP community nutrition program found that although women in program communities had higher levels of knowledge than women in nonprogram areas, the impact of the program on nutritional outcomes was small (IEG 2005; White and Masset 2007).[34] There are factors constraining women from acting that are not gender related (for example, resources, time), but the authors of one study point to evidence from a Demographic and Health Survey (DHS) that women are often not the main decision makers with respect to nutrition decisions in Bangladesh (IEG 2005).[35] In many cases, men do the shopping and mothers-in-law make meal decisions.

Differences in the age of the children studied are partly responsible for the variability in results.

If there truly are certain ages at which children are more susceptible to nutritional shocks and more likely to recover from them, then programs would be expected to have different impacts, depending on the age of the target group. The evaluations reviewed here did not consistently report results for similar age groups. The three evaluations of de-worming, for example, examined the impact on children 1–7 years old in Uganda, 2–6 years old in India, and 6–18 years old in Kenya (respectively: Alderman and others 2006a; Bobonis, Miguel, and Sharma 2006; Miguel and Kramer 2004). These results are not easily compared with findings on community-based nutrition programs, which measured impacts on children under 3 years (3 evaluations), under 5 years (1 evaluation), 6 months–2 years (2 evaluations), and 5–30 months (1 evaluation).

> Some of the variation in results is due to evaluation of impacts in different age groups.

Many of the studies measured impacts only on a relatively large age spread, such as 0–60 months, without reporting disaggregated results for children under 2 or 3 years old. This points to the possibility that some of the statistically insignificant findings for broad age groups might have yielded different findings had the age groups been disaggregated. For example, there was no program impact of the Uganda early child development program on WAZ of children aged 0–48 months, but when the author studied only children under 12 months of age, WAZ improved (Alderman 2007).

Although the age group of analysis is contributing to the variability in results in the aggregate, there is still variability in results among children of the same age. Comparing all studies that examined age groups under 36 months and controlling for the anthropometric outcome measure, evaluations even of the same intervention show inconsistent results, with some showing impacts and others none. The results and the age groups studied are sufficiently variable that this review could not confirm a pattern of higher program impact for children under three years of age, corresponding to the critical window of opportunity to prevent malnutrition (Agüero, Carter, and Woolard 2007; Allen and Gillespie 2001; World Bank 2006a).

> Short durations of exposure to the programs may explain low impacts in some cases.

Increased exposure raises impact.

Impacts are affected by duration of exposure to the program. Interventions that are implemented for a few months may not have a discernible effect on linear growth. Some of the reviewed evaluations mention short duration of exposure as a justification for lack of impact on stunting (for example, Bobonis, Miguel, and Sharma 2006; Kazianga, de Walque, and Alderman 2009; and Santos and others 2001).

Differences in duration of exposure can result in differences in magnitude and significance of impacts of the same program (Agüero, Carter, and Woolard 2007; Armecin and others 2006; Galasso and Yau 2006).

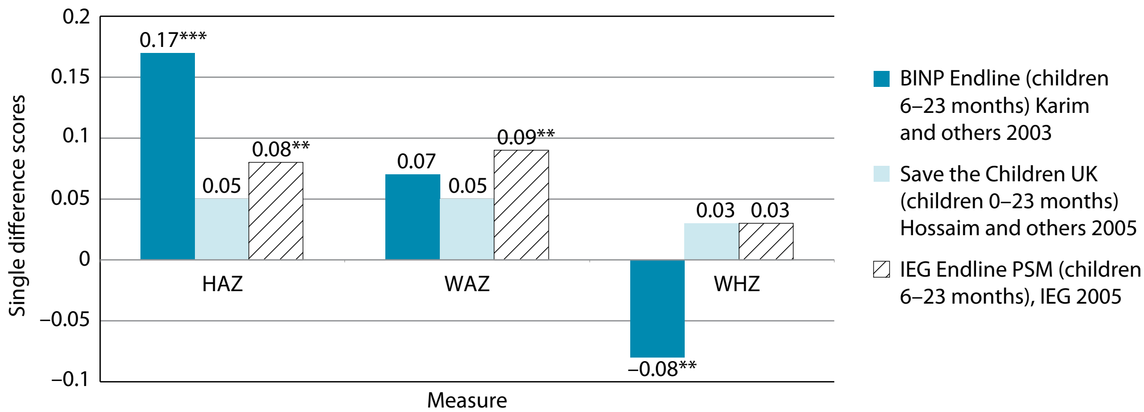
Evaluation methodologies can affect the results.

Studies that evaluated the same program using different methods arrive at different results. On the basis of experimental results of a nutrition education intervention in Peru, Penny and others (2005) report a significant difference in the WAZ of children aged 18 months in control and intervention areas. However, in a multivariate analysis of the same program, Waters and others (2006) show that the difference disappears when controls are included for selected socioeconomic characteristics.

> The impact of nutrition education in Peru depended on which estimation method was used.

The evaluations of the BINP community-based nutrition program in Bangladesh on nutritional status of children under two years old tell a similar story (figure 2.2). Early project monitoring data showed substantial reductions in malnutrition, especially in severe malnutrition, in project areas and convinced the World Bank and the government to scale up the intervention in the National Nutrition Project (Karim and others 2003). A subsequent evaluation of the program that compared program and nonprogram areas found no difference in stunting, underweight, or wasting between the program and nonprogram areas (Hossain and others 2005). However, it was unclear how well matched the program and nonprogram areas were in terms of their baseline characteristics before the program was launched. Using propensity score matching, IEG's reanalysis of the

FIGURE 2.2 Child Anthropometry Findings of Three Evaluations of the BINP



Source: IEG 2005, Tables G.17 and G.19.
Note: Significance levels: ** p ≤ .05; *** p ≤ .01. IEG endline PSM estimates are average treatment effects on the treated, one-to-one matching. BINP = Bangladesh Integrated Nutrition Project; HAZ = height-for-age z-score; WAZ = weight-for-age z-score; WHZ = weight-for-height z-score.

same data suggested that the project had a modest impact at best (IEG 2005; White and Masset 2007).

Conclusions

This chapter synthesizes evidence from 46 recent evaluations that analyzed the impact on child anthropometric outcomes of interventions implemented in 25 developing countries. More than half of the studies show impacts on at least one anthropometric indicator for some children. However, the lack of disaggregated results for common age groups makes it difficult to compare results across evaluations, and inadequate evidence on the causal chain and cost-effectiveness of the programs makes it difficult to synthesize the lessons.

The results are enormously variable, which is partly explained by context, the child's age, duration of exposure to the program, and the analytical methods used. Although there may be biological factors that justify early action, the evaluations of the programs reviewed here do not consistently show short-term impacts over the window of opportunity among the youngest children, during which time impacts are anticipated to be greatest.

Finally, most of the evaluations focused on average impacts; among the minority that measured the distribution of impacts there were differential impacts by socioeconomic status and mother's education. Only 1 in 8 of the evaluations addressed impacts by gender.

Chapter 3

Photo by Curt Carnemark, courtesy of the World Bank Photo Library.

EVALUATION HIGHLIGHTS

- Twelve nutrition impact evaluations evaluated interventions or programs in eight countries receiving World Bank support.
- Cash transfers, community nutrition, and early child development programs were evaluated.
- A large majority of evaluations used quasi-experimental methods.
- Evaluating large programs presented many challenges.
- The degree of implementation of the interventions was not well documented.
- Only half of the evaluations examined the heterogeneity of impacts; fewer documented costs.
- The impact evaluations in two of the eight countries plausibly had an impact on policy.

Evaluations of World Bank Nutrition Support

In an effort both to increase knowledge and to improve the effectiveness of programs, the World Bank has embarked on major initiatives to support rigorous impact evaluations, often embedded in World Bank projects.[1] IEG's recent evaluation of the Bank's support for health, nutrition, and population (HNP) found that though nearly a third of HNP projects called for impact evaluations or evaluation of pilot projects in their design, only about 1 in 20 actually conducted one (IEG 2009). Thus, a review of the characteristics, implementation experience, and ultimate impact of nutrition impact evaluations on policy is likely to lead to valuable insights on how to improve their effectiveness.

This chapter reviews the experience of the 12 evaluations that assessed World Bank–supported interventions to reduce malnutrition from among the 46 reviewed in chapter 2. Specifically, it reviews the characteristics of the programs evaluated, the challenges of designing and implementing impact evaluations of large government programs to reduce malnutrition, the evaluations' findings, the impact of the evaluations on programs and policy, and the lessons that can be drawn. The evidence is culled from a review of project documents, the evaluations, and interviews with project managers, evaluators, and country policy makers.[2]

Twelve evaluations measured the impact of Bank support on nutrition outcomes in eight countries.

The Programs Evaluated

Twelve of the 46 recent nutrition impact evaluations reviewed for this study could be linked to interventions supported by eight projects financed by the World Bank (table 3.1).

- Evaluations in Colombia and Ecuador examined the impact of **CCTs** and **UCTs**, respectively, on child nutritional and development outcomes.
- Evaluations in Bangladesh, Madagascar, and Senegal measured the impact of **community nutrition** interventions. These programs involved growth monitoring promotion for young children, nutrition education for the mothers (including breastfeeding messages), micronutrient supplements, and, in Bangladesh and Madagascar, food supplements for severely malnourished women or children. The services were delivered by community workers, supervised by NGOs.
- Evaluations in Bolivia, the Philippines, and Uganda measured the impact of **early child development** interventions on nutritional outcomes. The program in Bolivia consisted of informal, home-based day care that included nutrition supplements, stimulation, and access to health care. The early child development programs in the Philippines and Uganda had community-level workers providing nutrition services, in addition to early child education interventions. An ancillary impact evaluation embedded in the Uganda early child development evaluation assessed the impact of de-worming on the weight of preschool children.

Bank-supported cash transfers, community nutrition, and early child development programs were evaluated.

With only one exception, the World Bank–supported programs that were evaluated were large-scale government programs with multiple interventions and a very long causal chain that involved the compliance of implementers as well as beneficiaries to ensure effective implementation. Only the de-worming program for preschool children in Uganda had a relatively short results chain and comprised a single intervention implemented in a discrete region. All programs evaluated were implemented by developing country governments (national or local) or by NGOs on contract to government. This is in contrast with the larger body of nutrition impact evaluations reviewed by Bhutta and others (2008), most of which involved randomized controlled trails (RCTs) of discrete interventions with a short causal chain.

TABLE 3.1 Impact Evaluations of Programs and Interventions Supported by World Bank Projects

Country/ intervention	Project	World Bank support (total project cost)[a]	Program/nutrition intervention evaluated	Impact evaluations
Cash transfers				
Colombia	Human Capital Protection (2001–05)	$152.5 ($369.3)	*Familias en Acción* CCT: Subsidy of $15.38/month/mother of children 0–6, conditioned on growth control and development checkups every 2 months for children <1 year, three times a year for children <2 years, and twice a year for those 2–6. Subsidy of $4.61 for a child in primary school and $9.23 for secondary school, ages 8–17, conditioned on 80% attendance.[b] Targeted to the poorest quintile of the population in program municipalities.	Attanasio and others 2005
Ecuador	First Programmatic Human Development (2003)	$50.0 ($50.0)	*Bono de Desarrollo Humano* UCT: $15/month/mother, $11.5/month for senior and disabled for the lowest two poverty index quintiles[c] (amounts to a 10% increase in monthly income for average eligible family).	Paxson and Schady, forthcoming
Community nutrition				
Bangladesh	Integrated Nutrition (1995–2002)	$58.6 ($65.7)	Monthly growth monitoring and promotion for children <2 and pregnant or lactating women (PLW); supplementary feeding of malnourished PLW and severely malnourished and growth-faltering children <2; nutrition education for pregnant women, mothers of children <2, and adolescent girls at community nutrition centers.[d] Implemented by community nutrition promoters linked to NGOs in some *thanas*, by government in others.	White and Masset 2007/IEG 2005; Hossain and others 2005
Madagascar	Community Nutrition II (1998–2009)[e]	$46.7[f] ($67.9)	Community nutrition worker supervised by an NGO provides growth monitoring of children under 3, vitamin A to children and PLW, education and cooking demonstration, and food supplementation for severely malnourished children. Evaluation of program from (1) 1999–2002 and (2) 1997–2007.	(1) Galasso and Umapathi 2009; (2) Galasso and Yau 2006
Senegal	Nutrition Enhancement (2002–06)	$14.4 ($23.1)	Growth monitoring and promotion, behavior change communication through female CNAs for pregnant women and children <3, basic health services by CNAs (managed by NGOs) and district health staff (vitamin A, iron, de-worming, insecticide-treated bednets, sick child consultations), for children <5 and PLW. Evaluation of program 2004–06.	Alderman and others 2009; Linnemayr and Alderman 2008
Early child development				
Philippines	Early Childhood Development (1998–2006)	$17.3 ($49.6)	Enhanced early child development services to pregnant women and children <7. Integrated and multisectoral approach to delivering center-based and home-based interventions, linked by a child development worker who offers food and nutrition supplements, monitoring of health status, and parent education. Upgraded facilities.	Armecin and others 2006
Uganda	Nutrition and Early Childhood Development (1997–2005)	$35.1 ($40.6)	(1) Community-based growth monitoring and promotion carried out by childcare workers, and "child days" offering vitamin A, routine and catch-up immunization, and promotion of family care practices, children 0–6. (2) De-worming of children 1–7 years old, in addition to the standard community nutrition program. Evaluated from 2000 to 2003.	(1) Alderman 2007 (2) Alderman and others 2006
Bolivia	Integrated Child Development (1993–2004)	$25.6 ($60.6)	Nonformal, home-based day care centers (PIDIs) for children between the ages of 6 months and 6 years in poor families in 34 low-income urban areas. Services included food supplements for malnourished children (70%–100% of daily requirement), access to health care, and early child education. Women trained to offer childcare for up to 15 children in their homes, with a grant/loan of up to $500 to bring the home to minimum standards of safety and hygiene; two to three caregivers per home from the community trained by the project. Evaluated for 1996–98.[g]	Behrman, Cheng, and Todd 2004

Sources: Project Appraisal Documents and Implementation Completion and Results Reports.

Note: CCT = conditional cash transfer; CNA = community nutritiion aide; NGO = nongovernmental organization; PIDI = Integrated Child Development Project; PLW = pregnant or lactating women; UCT = unconditional cash transfer.

a. Expressed in $ millions. Final disbursements, except for Madagascar Community Nutrition II, which was still active as of October 2009.
b. 2002 data.
c. BDH was supposed to be conditioned on children between 0 and 5 years of age receiving bimonthly visits to health posts for growth and development checkups and immunizations and 90% attendance for children aged 6–15 years. However, these conditions were not enforced.
d. There were several other activities at the national level (information, education, and communication, vitamin A supplementation, and salt iodization, gardens, and poultry) that either never materialized, had weak capacity building compared with what was envisaged, or were implemented on a limited scale (Pelletier and others 2005).
e. Project still active.
f. Originally $26.7 million, plus two additional financings of $10 million each.
g. Beginning in 1997, the intervention underwent substantial changes to reduce the cost and increase the coverage.

Almost all the programs were large-scale government programs with many interventions and long results chains.

Three-quarters of the programs evaluated were completely new government programs. *Familias en Acción* in Colombia was a CCT program that had only been piloted in a few towns and was to be launched on a large scale. Ecuador's BDH was to be a better-targeted CCT, replacing an unconditional, poorly targeted program (*Bono Solidario)*.[3] The community nutrition programs in Senegal (PRN) and Madagascar (SEECALINE) had been previously piloted and were evaluated in the first major scale-up phase.

However, in Uganda there had been no pilot for the early child development program. It was evaluated in one region of the country, while the program was national in scope, targeted to the most malnourished areas. The PIDI child care program in Bolivia was also totally new, based on only a year's experience with pilot activities and modeled after a successful program in Colombia.[4]

In contrast, two of the programs were ongoing when evaluated. The early child development program in the Philippines aimed to improve ongoing services through better inputs and a multisectoral delivery mechanism that used a new type of community worker. The IEG-financed impact evaluation of the community nutrition activities of the BINP (White and Masset 2007) arose out of a need to reconcile conflicting findings of impact evaluations generated by the project and by the Save the Children Federation (Hossain and others 2005), neither of which had robust control groups.

Most of the impact evaluations were foreseen at project appraisal.

Most of the impact evaluations were foreshadowed in the Project Appraisal Document (PAD) as part of the project's monitoring and evaluation plan.[5] The PAD for the Bolivian early child development project, PIDI, defined the intervention group and two control groups; the impact evaluation of PIDI was part of the project's monitoring and evaluation component.[6] The Colombian Human Capital Protection Project PAD called for an evaluation with "a comparison group that will provide a counterfactual for what would have occurred had the Project not been implemented"; the evaluation was to be external. Preparation and implementation milestones of the impact evaluation of Ecuador's BDH were triggers for the approval of each of the three planned operations in the Programmatic Human Development Reform series.[7]

In contrast, the Senegal Nutrition Enhancement Project did not explicitly mention an impact evaluation above and beyond normal project monitoring and evaluation. However, one of the triggers for moving from Phase I to Phase II of the Adaptable Program Loan was an independent evaluation of Phase I. In Madagascar, the impact evaluation was not foreseen until well after the Community Nutrition II Project was approved. The BINP conducted an evaluation at the end of the project that was said to measure impact (Karim and others 2003), but the two external evaluations reviewed here (Hossain and others 2005; White and Masset 2007) were conducted after the project closed and were not foreseen.[8]

Most of the impact evaluations involved World Bank researchers.

Three-quarters of the evaluations, representing six of the eight projects, were led by or done in coordination with researchers in the Bank's Development Research Group. The three exceptions were the evaluations of BINP by Hossain and others (2005), sponsored by Save the Children Federation/UK (SCF), and by White and Masset (2007), sponsored by IEG; and the evaluation of *Familias en Acción* in Colombia, for which the government contracted with a consortium of research groups (Attanasio and others 2005).[9]

World Bank researchers were involved in conducting the evaluations of six of the eight projects.

World Bank research evaluators often participated in project preparation or supervision, but not directly in data collection. In Ecuador and Uganda, the researchers participated in project appraisal missions, and in all six countries they participated in supervision missions (financed through Bank operational budgets), either to supervise the implementation of the impact evaluation or the other project monitoring and evaluation activities.[10]

Other than influencing the timing of the rollout of the interventions in Ecuador and Senegal, the evaluation designs were not reported to have affected the design of the project or the intervention. In Colombia, Ecuador, and Senegal, household surveys were contracted out to private firms, some of which had experience in implementing the DHS. In Bolivia and Madagascar, the data were collected by national statistical offices. Only in the Philippines and Uganda were university research institutes directly responsible for data collection.[11]

Some of the evaluations were linked to program monitoring data.

The evaluations drew to varying degrees on program monitoring data. The cash transfer evaluations in Colombia and Ecuador used data from banking and administrative systems to verify the timing and amount of the transfers to beneficiaries. The evaluation of early child development in the Philippines used administrative data to document exactly when the improved services became available. The evaluation of the community nutrition intervention in Madagascar used aggregated program data by site on the percentage of children who were malnourished (as collected by the community nutrition worker).

The evaluations used program monitoring data to different degrees.

The two evaluations of BINP drew on program data for the analysis of supplemental feeding of severely malnourished or growth-faltering children, and the White and Masset evaluation (2007) used the project's midterm and endline household survey data. The evaluation of the impact of adding de-worming for preschool children to Uganda's early child development program relied on the program's child-weight monitoring data. However, the evaluations of early child development programs in Bolivia and Uganda and of community nutrition in Senegal reportedly did not link to any program monitoring data.

The evaluations were financed from diverse sources.

It is difficult to obtain exact information on the costs of most of the evaluations, because all but one (*Familias en Acción*, Colombia) received funding from multiple sources. Further, in some cases more than one evaluation was conducted using the same data set (for example, the BINP evaluations by Hossain and others [2005] and by White and Masset [2007]), or one of the evaluations piggybacked on the other (the de-worming and early child development evaluations in Uganda). One would have to allocate the costs of data collection across multiple studies. The time costs of the World Bank researchers and academic evaluators are not easily documented. However, it is possible to document the sources of funding for these impact evaluations (table 3.2).

They were financed by projects, lending operations, World Bank budget, and trust funds.

Governments financed at least part or most of the impact evaluations—usually data collection—through the lending operation, whereas the data analysis was often subsidized from other sources. Seven of the eight projects financed data collection and, in some cases, analysis of the data used for the 12 impact evaluations. The Senegal PRN project financed $700,000 for the first- and second-round surveys for the impact evaluation (World Bank 2007b). Only the evaluation of Colombia's CCT, *Familias en Acción*, was completely funded by the project, including data collection and analysis. In contrast, very little of the evaluation of Ecuador's BDH, an unconditional cash transfer program, was financed by the government.[12] The data used by Galasso and Yau (2006) in Madagascar were entirely from routine administrative sources and entailed no additional data collection expenditure.

The World Bank research budget supported evaluations in six of the eight countries.

Evaluations in six of the eight countries also received support from grants by the World Bank Research Committee for two research proposals for a total of $600,000.[13] Other sources of finance for either data collection or analysis included World Bank project supervision budget support

Photo by Yuri Kozyrev, courtesy of the World Bank Photo Library.

TABLE 3.2	Sources of Funding for Evaluations of the Impact of World Bank–Supported Programs on Nutrition Outcomes						
		Sources of funding					
Country	Project[a]	Project	World Bank supervision budget	World Bank researcher time	World Bank research support	Trust fund	Other
Cash transfers							
Colombia	Human Capital Protection/FA	√					
Ecuador	First Programmatic Human Development Reform/BDH		√	√	√	√[b]	
Community nutrition							
Bangladesh[c]	Integrated Nutrition/BINP	√				√[d]	√ IEG, SCF
Madagascar	Community Nutrition II/ SEECALINE[e]	√		√	√	√[f]	√ UNICEF
Senegal	Nutrition Enhancement/PRN[e]	√	√	√	√		
Early child development							
Bolivia	PIDI	√	√		√		√[g]
Philippines	Early Childhood	√	√	√	√		√[g]
Uganda	Nutrition and Early Childhood[e]	√	√	√	√		√[g]
Total		7	5	5	6	3	5

Sources: Interviews with task team leaders and evaluators, research committee funding proposals, and PADs.

a. BDH = *Bono de Desarrollo Humano;* BINP = Bangladesh Integrated Nutrition Project; FA = *Familias en Acción;* PIDI = *Proyecto Integral de Desarrollo Infantil;* PRN = *Programme de Renforcement de la Nutrition;* SEECALINE = *Projet de Surveillance et Éducation des Écoles et des Communautés en Matière d'Alimentation et de Nutrition Élargi.*

b. Japanese Policy and Human Resources Development Fund Grant, Spanish Impact Evaluation Fund.

c. Includes sources of funding for all three BINP evaluations—by the project team, by SCF, and by IEG.

d. Department for International Development partnership, Danish Trust Fund.

e. Includes funding sources for more than evaluation of the program.

f. Bank Netherlands Partnership Program Trust Fund.

g. Co-investigators brought funding from additional sources.

(Ecuador, Madagascar, the Philippines, Senegal, Uganda), trust funds (Bangladesh, Ecuador, Madagascar),[14] IEG budget (Bangladesh),[15] and research funds from academic co-investigators.

The Design and Implementation of the Evaluations

Most evaluations used quasi-experimental designs.

Few of the nutrition impact evaluations attempted to randomize the assignment of the program; those that did so randomized assignments at the community, not the individual, level. Only three of the evaluations attempted to randomly assign the program, and of these only one was able to maintain a relatively clean design during project implementation. The rollout of the BDH unconditional cash transfer was randomized such that the communities receiving the intervention in future years could serve as the control group for the communities that received the intervention at the start (Paxson and Schady, forthcoming). The random assignment to the rollout was maintained.

However, in Uganda and Senegal, the randomized assignments did not go according to plan. A de-worming intervention for young children in Uganda was randomly assigned among areas already receiving an early child development intervention, but some households in the control group nevertheless increased purchase of de-worming medicine on their own (Alderman and others 2006).

The attempt to use a randomized program rollout to provide for treatment and control groups in Senegal for the PRN was foiled when the NGOs responsible for implementing the program did not adhere to the plan—postponing its launch in some of the treatment areas and implementing it earlier than planned in some control areas (Alderman and others 2009; Linnemayr and Alderman 2008).[16] Nevertheless, the evaluators in both of these cases were able to use the randomized assignment as an instrumental variable to predict treatment, purging the impact estimates of self-selection bias.

Three of the impact evaluations had randomized designs, but the designs for two were not fully realized.

For either political or practical reasons, most of the evaluations used quasi-experimental methods for estimating program impact. Policy makers in Colombia, for example, were unwilling to embrace randomized rollout of interventions at a time of political crisis. In the evaluation of *Familias en Acción*, Attanasio and others (2005) compared randomly selected treatment municipalities with matched control municipalities on the basis of geographic region, education and health infrastructure, population, and other characteristics.[17] They estimated the impact based on the difference-in-difference between treatment and control areas over time. However, there were still fundamental differences between the baseline treatment and control areas that led to the use of propensity score matching to generate a control group.[18] In Uganda, Alderman (2007) compared project areas with controls that were nonproject subcounties adjacent to each subcounty in the study; the areas were found to be sufficiently similar in characteristics to simply compare the mean effects between the treatment and controls.

Most of the evaluations of the nutrition impact of Bank-supported programs had a quasi-experimental design.

The evaluations that drew on existing data sets or programs already under way did not have the option of a prospective experimental design. For example, the BINP in Bangladesh was ongoing when evaluated by two sets of researchers, which led them to choose matching methods. The project and nonproject comparison areas used by Hossain and others (2005) were not good matches; White and Masset (2007) used the BINP project survey data for the treatment areas but used PSM to generate a control group using a third, nonproject data set.

The evaluation design for the PIDI program in Bolivia called for comparing a random sample of program participants with two matched comparison groups—one consisting of households and children nationwide with characteristics similar to those of the treatment group and the other of households in the same neighborhood that did not enroll their children in the program (Behrman, Cheng, and Todd 2004). These groups were not found to be sufficiently similar; consequently, the authors used matching methods to control for selectivity into the program. *In fact, 7 of the 12 evaluations employed PSM, either because they had no control or comparison group or because the selected control groups were found to be inadequate (*Armecin and others 2006; Attanasio and others 2005; Behrman, Cheng, and Todd 2004; Galasso and Umapathi 2009; Galasso and Yau 2006; Linnemayr and Alderman 2008; White and Masset 2007).[19]

Evaluations in Bolivia and Madagascar compared cohorts exposed to the programs for different amounts of time.

Evaluations in Bolivia and Madagascar estimated marginal impacts of program exposure by comparing cohorts of participants who had been in the program for different amounts of time with those who had only recently joined. In Bolivia, children enrolled in PIDI for two months or more (up to more than 25 months) were compared with children enrolled for a month or less. In Madagascar, communities that had participated for two years were compared to matched communities that had participated for one year, and both were compared with communities that had just enrolled. The evaluation used regularly collected administrative data of the community nutrition program, supplemented in later phases by household surveys (box 3.1). One of the advantages of this approach is that examining the effects of additional exposure does not require a control group.[20]

People in control groups spontaneously adopted the same activities as those assigned to the treatment groups in the Philippines and Uganda.

Crossover effects were experienced in evaluations with comparison groups as well as in those with control groups to which the intervention was assigned on a random basis. Parents of about a third of the Ugandan children in the control group got their children de-wormed (Alderman and others 2006). In the Philippines, nonproject areas spontaneously adopted some of the activities of the early child development program being evaluated (Armecin and others 2006). In both cases, these crossover effects resulted in muting the difference between the treatment and control or comparison areas.

BOX 3.1 **Measuring the Impact of Additional Exposure to a Community Nutrition Program Using Program Data in Madagascar**

The second Community Nutrition Project in Madagascar supported a community-based nutrition program implemented by community nutrition workers (CNWs) supported by NGOs. According to the PAD, the CNW is elected by the community, trained, and receives an annual salary of about $350–$400. The community identifies a nutrition center, can get a grant of up to $200 to furnish it, and receives basic weighing and measuring equipment. The centers are to cover a population of 2,000 with the capability of covering 226 children within 5 kilometers. Social workers are also recruited by local NGOs. The CNW conducts a census of all children under three years of age at the outset and annually thereafter. The CNW weighs all children under three monthly and gives the mothers nutrition education and a cooking demonstration. Malnourished children get food supplements and are monitored every two weeks. Children who weigh in at < –3 SD WHZ are sent to the health system for rehabilitation. Vitamin A supplements are given once a year to children under 24 months, twice a year to children 24–36 months, and to lactating women within six months of delivery.

The impact evaluation used aggregated routine monitoring data from 1999 to 2002 from four main provinces, from about 3,600 sites and about a quarter of all communities in the country. The authors used the time delay involved in the rollout of the program to compare participating communities with one or two years of intervention with communities just starting. Because the phase-in began with the most severely affected communities that also had NGOs, the authors used PSM to adjust for selection bias. The evaluation found that two years' exposure to the program reduced the percentage of children under three years or age who were underweight by 7–9 percentage points, from an initial level of 46 percent.

Sources: Galasso and Yau 2006; World Bank 1998.

The evaluations measured short-run impacts.

Most of the evaluations assessed impact over a relatively short period following the launch of the intervention. Two-thirds of the evaluations measured impact after no more than three years of implementation and, in 7 of the 12 cases, two years or less.

The quality of the service may improve over time following a learning curve, and longer exposure may independently affect the impact if there is a dose-response relationship. For these two reasons, somewhat less impact may be expected for certain interventions (for example, for an intervention to affect chronic malnutrition) over a relatively short implementation period. Failure to control for the actual launch date can result in an underestimate of the impact or to a finding of no impact at all.

> The evaluations measured short-run impacts on malnutrition, generally within two years of program start-up.

The initial findings of the Philippines early child development impact evaluation found little or no impact; not until the researchers went through administrative records to pinpoint when services became available for each community did significant results appear (Armecin and others 2006). The two evaluations of BINP may suffer from this problem; only the rough starting date—about 1996—of the overall program is known, yet the program was supposed to be phased in over a five-year period. This was apparently not controlled for in either evaluation.

Evaluation of large-scale government programs presented challenges.

The evaluations faced major challenges because of delays in project launch, disruptions in delivery, political pressures not to follow the plan, and disruptions caused by political pressure, natural disasters, and other breakdowns in program implementation.

Delay in launching the intervention. In Uganda the early child development program baseline survey was done in January–March 2000, but the growth-promotion intervention did not begin until late 2001 and the community nutrition grants started in 2002. As a result, the intervention had been operational for only a little more than a year by the time of the endline survey in January–March 2003.

> Delays in project effectiveness delayed baseline surveys in Bolivia and the Philippines.

In Bolivia the baseline data collection was postponed two years because of a delay in project effectiveness. Partly because of the extensive delays in launching the early child development project in the Philippines, the results of the first round of evaluation found very little impact. This led the researchers to seek from administrative records exactly when the intervention had been launched in each site, to be assured that the project areas were, in fact, exposed. The delays, however, were not always detrimental: in Senegal, the delay in project implementation allowed time to design the impact evaluation.

Political pressures not to follow the plan. The evaluation of BDH in Ecuador was supposed to have three arms—a CCT, a UCT, and a control group. However, the government never enforced the conditionality, so for all intents and purposes, it was an unconditional transfer and there were twice as many treatment households as there were controls.

In Colombia, there was an election and political change shortly after the researchers were awarded the contract to collect the baseline data. This created pressure to scale up the program before the baseline data could be collected. At the time that the baseline data were being collected, in 2002, some towns were already participating in the program. In Uganda, pressure from parliament led to the expansion of the project to more districts than planned without increasing the budget. Although this did not expand the scope of the impact evaluation, it reduced the resources for implementing the project, and the project ran out of money before many activities could be completed.

In Colombia the program was expanded before the baseline could be implemented.

In Senegal, during the delay in project implementation, NGOs conducted social mobilization to prepare and organize the communities destined to participate in the project. This made it difficult, once the evaluation design was finalized, for the researchers to explain to some communities that in fact the services would be delayed a year or two because of the need to randomize the rollout. In fact, the NGOs in charge of implementing the intervention did not respect the randomization of communities, electing to launch the intervention in some phase 2 areas and delay it in the phase 1 areas. As a result, 30 percent of the villages that had been randomly selected to get the intervention in the first round did not get it, and eight of the control villages in the first round did (Alderman and others 2009).

NGOs in Senegal mobilized communities before the impact evaluation design was finalized, making it difficult to respect the randomization plan.

Disruption in service delivery caused by changes in the political context, natural disasters, or breakdowns in program implementation. In Madagascar the SEECALINE project was amended five times, with two additional financings, in response to cyclone damage in 2000 and 2004 and to political turmoil in 2002. The first of the restructurings added rural areas of 16 more districts to the 52 districts already targeted and urban areas of 6 districts, representing 550 more sites. The 2006 amendment expanded the program to include children under five in selected communities in all 110 districts of the country. Beyond this, there were regular disruptions in the availability of food for the take-home rations that were to be issued to children who did not gain weight for two months.

In Bolivia, within two months of approval of the project in 1993, a new administration took office that had concerns about the scale and financing of PIDI. In 1994, the Decentralization and Popular Participation Laws were enacted, which made municipalities and departments responsible for social service investment decisions, and at the end of 1995, the implementing agency was dissolved and the project was assigned to the Social Investment Fund. In the Philippines early child development project, there were several changes in the Project Management Unit. Following each change, the researchers had to rebuild support for the impact evaluation.

Findings

Three-quarters of the 12 impact evaluations found a positive impact on anthropometric outcomes of children in at least one age group, although the magnitude was in some cases not large or the impact applied to a narrow age group.[21] The evaluations are notable not only for the variability in their findings (discussed below) but also for the extent to which the complex results chain was documented, so as to put forward a plausible story of causation and to understand the extent to which the interventions were actually implemented. When implementation is spotty, it can be as if there is no intervention at all. The anthropometric impacts and the extent to which the evaluations documented program outputs and intermediate behavioral outcomes are summarized in table 3.3.

Three-quarters of the evaluations found program impacts, but little is known about what part of the intervention worked.

Average impacts for similar interventions were variable; links to the underlying causal chain were weakly documented.

Cash transfers. Colombia's *Familias en Acción*, a CCT program, would be expected to affect nutrition status through the additional income of the cash transfer and the conditionality on use of health and education services. However, in Ecuador's BDH, an unconditional cash transfer program, only the income effect would be operating. The evaluations

TABLE 3.3 Nutrition Impact Evaluations and the Results Chain for World Bank Projects

Program type	Country	Evaluation	Evaluation period	Program output data analyzed?	Intermediate outcomes analyzed?	Was there an anthropometric impact?	Heterogeneity of impacts analyzed?
Conditional cash transfer	Colombia	Attanasio and others 2005	2002–06	Yes. Administrative data on payments, health, and education service data.	Yes. Diphtheria, pertussis, and tetanus vaccination rate; reported food intake; participation in growth monitoring.	Yes. HAZ (especially for children <24 months) and newborn weight.	No
Unconditional cash transfer	Ecuador	Paxson and Schady, forthcoming	October 2003/ September 2004–September 2005/ January 2006	Yes. Bank records of transfers and when started.	Yes. Participation rate; use of health clinics for growth monitoring; sought treatment for helminth infections.	No. HAZ (ages 3–7).	Yes. Household poverty; gender.[a]
Community nutrition, including food supplements	Bangladesh	Hossain and others 2005	1996–2002[b]	Yes. Children receiving food; effectiveness among those enrolled; food leakage; food substitution; village health worker quality.	Yes. Mother's nutrition knowledge and reported practice.	No. WAZ, HAZ, WHZ (ages 6–23 months).	No
		White and Masset 2007/ IEG 2005	November/ December 1998–January/ March 2003	Yes. Receipt of counseling; receipt of food; targeting of food; duration of food.	Yes. Participation rate for weighing; nutrition knowledge; practice.	Yes. WAZ and HAZ (age 6–23 months), but small in magnitude.	Yes. Household assets and mother's education.
	Madagascar	Galasso and Yau 2006	1999–2002	Yes. Characteristics of the NGOs.	Yes. Registration rate.	Yes. Underweight (age <3 years), relatively large in magnitude.	Yes. Community poverty; cyclone-prone areas; length of lean season; access to safe water.
		Galasso and Umapathi 2009[c]	1997–2007	Yes. Receipt of vitamin A and message; tetanus injection during pregnancy; assisted delivery; possession of health card; receipt of nutritional counseling.	Yes. Breastfeeding; feeding practices; hygiene practices; diarrhea.	Yes. WAZ, underweight, HAZ, stunting, relatively large (age < 5).	Yes. Mother's education; low-poverty areas; proximity to road, hospital; access to safe water, electricity.

TABLE 3.3 *(continued)*

Program type	Country	Evaluation	Evaluation period	Program output data analyzed?	Intermediate outcomes analyzed?	Was there an anthropometric impact?	Heterogeneity of impacts analyzed?
Community nutrition, without food supplements	Senegal	Alderman and others 2009	2004–06	No	Yes. Receipt of iron supplements; malaria pills by mothers; receipt of vitamin A; de-worming; ownership of bednets.	Yes. Underweight (age <3).	No
		Linnemayr and Alderman 2008	2004–06	No	Yes. Health inputs; nutrition knowledge of mother; breastfeeding practices.	Yes. WAZ (age <3).	Yes. Villages with seasonal roads; villages with lower average wealth at baseline
Early child development, with food	Bolivia	Behrman, Cheng, and Todd 2004	1996–98	No	No	No. Weight percentile and height percentile (ages 6 months to 6 years).	No
	Philippines	Armecin and others 2006	1996–98	Yes. Early child development worker training and functions; feeding programs; parent education seminars; home-based day care; exact onset of program.	No	Yes. WHZ and wasting. Mixed results–HAZ and stunting (age <7).	No
Early child development, without food	Uganda	Alderman 2007	2000–03	No	Yes. Breastfeeding and weaning practices; reported foods fed to children.	Yes. WAZ among those <12 months.	Education; imputed expenditure
De-worming	Uganda	Alderman and others 2006	2000–03	Yes. Number of child health days; treatment intervals.	Yes. Uptake.	Yes. Weight (age 1–7 years).	No

Source: IEG analysis.
Note: HAZ = height-for-age z-score; WAZ = weight-for-age z-score; WHZ = weight-for-height z-score.
a. The interactions for heterogeneity are not for HAZ individually, but rather a synthetic variable for physical development.
b. The authors assumed that the intervention began in 1996, the year the project was approved. However, implementation was supposed to be phased, and it is not clear when the intervention actually became available to the survey villages. Thus, the exposure may be significantly less than six years.
c. The outputs and intermediate outcomes are presented in Galasso and Umapathi 2009, a working paper that was revised for publication, from which this information was dropped.

of both programs document the disbursement of the transfers and changes in intermediate outcomes that would be consistent with improved nutrition outcomes—an increase in vaccination rates and reported food intakes in Colombia, treatment for helminth infections in Ecuador, and participation in growth monitoring in both countries.

In Colombia there was an impact on HAZ for children younger than two but not of children two to four or older. There was no impact of Ecuador's *Bono de Desarrollo Humano* on HAZ, but the children studied in that evaluation were between three and seven years of age. Thus, the finding of no impact for children over two is consistent across the two programs; it cannot be compared for younger children.

Another factor contributing to different findings could be that the transfers had been in place in BDH for two years or less, half the exposure of the *Familias en Acción* at the time of the evaluation. Different access to health care in the two countries could also have played a role, though that information was not presented.

Community-based nutrition programs. The six evaluations of community-based nutrition programs in Bangladesh, Madagascar, and Senegal generally found positive effects on weight and, when measured, height, though the size of the impact varied and many of the evaluations suffered from a lack of information on the extent to which the interventions were implemented.

The impact of BINP on WAZ and HAZ was small, even though the mothers' knowledge improved. It is not clear why. Many possible implementation factors could have been responsible; for example, the performance of the community nutrition promoters (CNPs) and the large number of people that each CNP was supposed to serve (more than 1,000). However, the evaluation did not explore this issue. It is also not known how long each of the communities was actually exposed, a factor that can result in underestimating impact.[22]

Photo by Curt Carnemark, courtesy of the World Bank Photo Library.

Food supplements for women or children typically account for a very large share of the costs of nutrition programs and are logistically difficult, yet their effectiveness was assessed only in the two BINP evaluations, even though food was an element in the programs for half of the eight countries.[23] White and Masset (2007), using data collected from health centers by the SCF authors (Hossain and others 2005), found important targeting problems. Only 16 percent of children receiving the food should not have received it, whereas more than two-thirds of the children who were eligible (that is, those with severe malnutrition or growth faltering) were not fed. Among those receiving food, only a quarter received the supplements for the recommended three months. More than 40 percent of the children who were receiving supplements were not malnourished, but were receiving them because their growth was faltering. The authors note, however, that growth faltering is normal.

An evaluation of Bangladesh's BINP was the only one that assessed the impact of supplemental food for malnourished children.

Community nutrition programs in Madagascar and Senegal had positive impacts on WAZ or underweight, primarily for children under three. In the case of Senegal (a nutrition program that does not dispense food), Alderman and others (2009) and Linnemayr and Alderman (2008) track important intermediate outcomes to explain those improvements—receipt of iron supplements and malaria pills by the mothers, receipt of vitamin A, de-worming, ownership of bednets, and breastfeeding. Yet neither of the evaluations for Senegal documents the extent to which the interventions were actually implemented. One of the evaluations of Madagascar's SEECALINE program documents changes in intermediate outcomes that are consistent with improved nutrition found in the evaluation—breastfeeding, hygiene, and feeding practices (Galasso and Umapathi 2009).

Early child development. The results of the evaluations of the early child development programs were likewise variable. Bolivia's PIDI, a nonformal, home-based day care program, had no effect on any anthropometric indicators for children six months to six years old, despite the fact that the program provided meals to the children amounting to 70 percent to 100 percent of their daily needs. There were significant program impacts on weight and wasting in the Philippines and Uganda, although only for children less than one year of age in Uganda,[24] and mixed effects on HAZ and stunting among children under seven, depending on the age group, in the Philippines.

Despite the large number of activities embedded in these programs—including growth monitoring and food supplements found in the community nutrition programs—the results chain of program outputs and intermediate outcomes for these three evaluations is weak.

The greatest challenge was for the evaluation in the Philippines of the improvement and reorientation of an existing early child development program. The research teams had to go from center to center to assemble the necessary administrative data documenting exactly when the intervention began. The evaluation shows convincingly that in the program areas the number of trained workers, feeding programs, day care centers, and other activities increased relative to the control areas. Even then, there is little evidence provided to demonstrate how well the services were delivered, and no information was presented on intermediate outcomes that might logically be linked to the nutritional outcomes observed.

> The impact of Uganda's early child development program is supported by changes in breastfeeding and weaning practices.

In contrast, the evaluation of early child development in Uganda presents no evidence on program outputs but does document changes in breastfeeding and weaning practices and in the foods reportedly fed to children (Alderman 2007). The evaluation of Bolivia's PIDI program, which found no anthropometric impacts, provides no information on either program outputs or intermediate behavioral outcomes that might explain this result (Behrman, Cheng, and Todd 2004).

> The PIDI early child program in Bolivia had no impact on height or weight, even though the children were fed.

De-worming. The single study that tested the impact of adding de-worming to the ongoing early child development intervention in Uganda found weight gains among preschool-age children (1–7 years) (Alderman and others 2006). The results chain for this particular intervention was short. The evaluation used administrative records to document each participating child's weight gain and the receipt of the de-worming drugs. There are few intermediate behaviors to document. However, it should be noted that the results are likely underestimates, as a sizable share of the parents in the control area spontaneously increased their purchase of de-worming medicine for their children. Further, the evaluation was launched in the region with the highest worm load, and both the treatment and control areas had access to the early child development intervention.

Only half of the evaluations documented heterogeneity in impacts.

Only half of the impact evaluations explored the distribution of impacts across individuals or communities.[25] The coverage of heterogeneity and the variables considered by each study are presented in the last column of table 3.3.

Poverty. Six of the 12 evaluations assessed whether poorer households or communities benefited more than the nonpoor. In Ecuador, the impacts were larger among the lowest quartile of eligible families (Paxson and Schady, forthcoming).[26] In Madagascar, the SEECALINE program, which was targeted to the poorest and most malnourished areas, had the largest impact on all four anthropometric outcomes in the better-off communities; in the communities with the highest poverty rates, only children of the most educated mothers had better anthropometric outcomes (Galasso and Umapathi 2009). In contrast, the sites with the highest poverty rates had higher returns to program exposure over two years (Galasso and Yao 2006). However, in Bangladesh, Senegal, and Uganda, there was no difference in impact in less wealthy households (IEG 2005), in poorer communities (Linnemayr and Alderman 2008), or in households with lower imputed expenditures (Alderman 2007), respectively.

> In Ecuador the benefits were greatest for the lowest income families, whereas in Madagascar children in better-off communities in the targeted poor areas benefited the most.

Mother's education and child's gender. Three of the evaluations assessed whether the impacts were greater for children of more educated mothers than for children of less educated mothers. In Madagascar, results suggested that children of educated mothers benefited more from the interventions (Galasso and Umapathi 2009); the impact of neither the Bangladesh community nutrition program nor the Uganda early child development program varied with mother's education (IEG 2005; Alderman 2007). Only one of the evaluations examined the impact according to the child's gender, finding that impacts were greater for girls than for boys (Paxson and Schady, forthcoming).

> In Madagascar, program impact was greater in communities with roads . . .

Availability of public services. Surprisingly, only three of the evaluations examined the relation between the program's impact and the availability of public services.[27] Even women with better knowledge of good child nutrition practices may be limited in their ability to act on this knowledge if they lack access to complementary services such as health care or to markets.

Galasso and Umapathi (2009) found that the impact of the Madagascar SEECALINE community nutrition program on all of the anthropometric outcomes was greater with proximity to a road or hospital, and that the WAZ impact was greater with access to a safe water source. However, the other evaluation in Madagascar, which used aggregated data across sites, found no difference in the returns to program exposure for communities with better access to safe water (Galasso and Yau 2006).

In contrast, the Senegal PRN community nutrition program had greater impact in more isolated villages not served by all-weather roads. That implies that the services of the nutrition worker may have been substituting for services outside the villages (Linnemayr and Alderman 2008).

. . . but in Senegal community workers substituted for the availability of services.

Program costs and cost-effectiveness were rarely assessed.

The impact evaluations rarely remarked on the program costs per beneficiary or conducted cost-benefit or cost-effectiveness analyses. In only three cases were costs presented in the published evaluations (or their antecedents), and in a fourth case (Madagascar), the analysis was done informally for the government based on the impact evaluation, but was not published.

The cost of the Bolivia early child development program was estimated by various sources to be as high as $43/month and as low as $22/month per child enrolled (Behrman, Cheng, and Todd 2004). Either cost clearly would be unsustainable for large numbers of children in Bolivia, with a gross domestic product/capita at that time of $800.[28] Nevertheless, the cost-benefit analysis done by the authors suggests a benefit-cost ratio (under varying assumptions and discount rates) between 1.37 and 3.66. This is based on the extrapolation of future benefits for the nonanthropometric impacts, however, as the study found no impact on HAZ or WAZ.[29] From the perspective of the actual nutrition outcomes, the benefit-cost ratio would be zero.

In unpublished calculations for government, the lead author for the two Madagascar evaluations calculated the unit cost of the SEECALINE program to be on the order of $7/child/year and the cost of preventing one child from being stunted as $219/child/year (Emanuela Galasso, personal communication). After discounting the benefits with various assumptions and discount rates, the benefit-cost ratio was estimated to be between 1.7 and 4.5.

Costs and cost-effectiveness of the programs were rarely assessed, and cost-benefit analyses were rarely performed.

Cost-effectiveness, in contrast to cost-benefit, can be more easily calculated in the context of an impact evaluation based on local data, actual implementation costs, and effects. Because the impact of the Bangladesh BINP as implemented is found to be so small, the cost to achieve a given outcome is high. The cost of preventing a child from being underweight was calculated to be $187–$333 per year, and for stunting $241–$490 annually, with an estimated cost per life saved ranging from $2,328 to $4,095 (IEG 2005).[30]

The marginal cost of adding de-worming medicine to the (then ongoing) early child development program in Uganda was calculated. Because the program was already distributing vitamin A to the children, only the marginal cost of $0.42 was included for twice-yearly de-worming treatment that would result in a 10 percent increase in weight gain (or half that amount for once-a-year de-worming) (Alderman and others 2006).

The Impact of the Evaluations

Is there any evidence that the findings of these 12 impact evaluations were used? This section pulls together evidence of the use of the data and other impacts from these evaluations based on a review of the projects' Implementation Completion and Results Reports (ICRs), the PADs of follow-on projects, any impacts of the findings mentioned in the evaluation reports, and interviews with key informants for each project—the World Bank project leaders, the evaluators, and at least one policy maker from six of the eight countries.[31]

Because it was not possible to conduct country visits, these findings should be considered partial and suggestive. Nonetheless, the findings across the documents and individuals consulted for each project were generally consistent. Table 3.4 summarizes evidence on the impact of the evaluations.

The impact evaluations plausibly had an impact on policy in two of the eight countries. In both countries, the intervention had a positive effect on child anthropometric outcomes.

In Colombia, a new political administration came into power in 2002, only a year after the project was approved. There was reportedly great concern at that time about the severe fiscal situation that affected all government programs and the high cost of the impact evaluation. However,

TABLE 3.4	**Summary of the Impact of the Nutrition Impact Evaluations**				
Country—intervention[a]	**Results reported in ICR?**	**Impact evaluation found nutrition impact?**	**Was there a follow-on project?**	**Were the results in the PAD?**	**Reported policy or program impact of the evaluation**
Cash transfers					
Colombia—*Familias en Acción*	Yes	Yes	Yes	Yes[b]	Yes. Generated political support to continue funding when new president came into power and to scale up; the evaluation also "contributed to defining the larger social protection and evaluation agenda in the country" (World Bank 2006, p. 11). Findings supported dropping the restriction that children born since program launch be excluded and including children also enrolled in *Hogares Communitarios.*
Ecuador BDH	No	No	Yes	No[c]	No. Government did not add conditionality and did not drop the next-to-poorest quintile, even though there were no benefits of targeting them. However, the evaluation greatly raised capacity in the ministry for conducting impact evaluations. Ecuador is pursuing impact evaluations of other programs.
Community nutrition					
Bangladesh BINP	No	Small	Yes[d]	No	No. Respondents report that the program has not changed. However, one respondent remarked that the Bank is paying more attention to the quality of service delivery as a result of the two evaluations.
Madagascar SEECALINE	[e]	Yes	[e]	[e]	Unclear. The project was expanded; the prime minister wrote a letter to *The Lancet,* along with the prime minister of Senegal. However, it appears that the program was politically popular even without the evaluation, so it is unclear whether it was really the impact evaluation that changed things.
Senegal PRN	No	Yes	Yes	No	Unclear. The program was scaled up in the second operation, which was the second phase of an Adaptable Program Loan; however, the results were not available at the time that decision was made. The evaluation may have been reaffirming.
Early child development					
Bolivia PIDI	Yes	No	No	n/a	No. The model evaluated was excessively expensive and subsequently adapted to a model quite different from the one evaluated. "All activities were ended as of December 2003 and none . . . were included or absorbed by other ongoing programs." "The family/home-based day care centers . . . have practically disappeared and most have been converted into community centers. Yet . . . they still have a high cost compared to other similar programs" (World Bank 2004, p. 30).
Philippines ECD	No[f]	Yes	No	n/a	Yes. Was reportedly used to justify expanding program innovations. Strong ownership of the impact evaluation; the ECD head presented results at the 2004 World Bank Conference on Scaling Up Poverty Reduction in Shanghai, China. However, the ECD program had strong support even before the evaluation showed some impacts. Possibly reaffirmed existing support.
Uganda ECD	Yes	Yes	No	n/a	No. Community nutrition has been dropped from the program, although child days have continued. (The idea of child days was mentioned as attributed to UNICEF.) None of the ministries, especially the Ministry of Health, ever owned the project.
De-worming					
Uganda ECD	Yes[g]	Yes	No	n/a	Unclear. The evaluation implies that the government expanded the de-worming policy following release of results, but others report that the decision to expand de-worming to preschool children had already been made. The evaluation may have influenced other African countries.

Source: IEG analysis.
Note: ICR = Implementation Completion and Results Report; PAD = Project Appraisal Document.
a. BDH = *Bono de Desarrollo Humano;* BINP = Bangladesh Integrated Nutrition Project; ECD = early child development program; PIDI = *Proyecto Integral de Desarrollo Infantil);* PRN = *Programme de Renforcement de la Nutrition;* SEECALINE = *Projet de Surveillance et Éducation des Écoles et des Communautés en Matière d'Alimentation et de Nutrition Élargi.*
b. There were two follow-on projects, both of which mentioned the results—the Social Safety Net Project (2005) and the Second Phase of the Program of Conditional Transfers—*Familias en Acción* (2008).
c. The PAD for the follow-on project discusses at great length the results for the impact evaluation on education, but not the results (or lack of results) on health and nutrition. The follow-on project was canceled following a change in government.
d. The follow-on project that scaled up BINP, the National Nutrition Project, was launched following the positive results reported for the BINP midterm review and before either of the evaluations (Hossain and others 2005; White and Masset 2007) was published.
e. This project was scheduled to close in December 2009. There was not yet an ICR at the time of this review. Whether there will be a follow-on project is not known.
f. The evaluation is mentioned and the trends in the treatment and control areas are charted, but the final evaluation results, as put forth in Armecin and others (2006), are not mentioned in the ICR.
g. The results of the de-worming are inaccurately conveyed in the ICR, which says that the largest impact (a 10% increase in weight) was among the youngest children (that is, those under 12 months). The magnitude is correct, but it was for children aged one to seven years; infants were not given de-worming medicine.

the results of the first wave of the evaluation of *Familias en Acción*, which became available shortly thereafter, showed impacts on schooling, health, labor supply, and consumption.[32] The government not only expanded the program to new areas and broadened the eligibility to additional children within the original areas but also embraced a program of rigorous impact evaluation more generally in developing its social safety net program.

The impact evaluations of Bank support plausibly had an impact on policy in two of the eight countries.

World Bank support was enlisted for two follow-on safety net projects, including additional financing for *Familias*. The PAD for one of two follow-on projects (Social Safety Net, approved in 2005) notes that "the program credibility has . . . been fostered by the very positive results of the conditional cash transfer evaluation that has been continuously disseminated since the early stages of program implementation" (World Bank 2005, p. 15). The full results are cited in the rationale for the Second Phase of the Program of Conditional Cash Transfers/*Familias en Acción* (approved in 2008, two years after the last round of data collection).

Early results from the Colombia's *Familias en Acción* helped convince a new administration not to cancel it.

The complete findings of the impact evaluation of the Philippines early child development project were not available at the close of the project; as the ICR was being written, only the trends in the project and nonproject areas were cited. There was already strong political commitment for the ongoing early child development program even as the program upgrades were introduced. Reportedly, since the project closed, many of the innovations have been incorporated more widely into the program. It is difficult to tell in this instance whether the evaluation merely reaffirmed the wisdom of something that government was already set to do or whether it had a role in the decision to expand the innovations.

In Madagascar, Senegal, and Uganda, evaluations found positive impacts on nutrition outcomes, but it was unclear whether subsequent program decisions were due to the evaluations. In Madagascar, following dissemination of the results of the evaluation of SEECALINE, the prime minister wrote a letter to *The Lancet* (cosigned by the prime minister of Senegal) extolling the positive impacts of community nutrition programs (Sall and Sylla 2005). The SEECALINE program was expanded multiple times over its more than decade-long lifetime, and the project supporting it still had not closed as of October 2009.

There have been changes in government following a period of unrest; however, elements of the community nutrition activities have been incorporated into the new National Nutrition Program. The program was politically popular, even before the evaluation, so it is unclear whether the evidence from the evaluation contributed to its expansion.

In Madagascar and Senegal, positive impacts may have helped maintain support for the programs.

The Bank's support for Senegal's PRN was packaged as part of a multiple-phased Adaptable Program Loan. The evaluation found evidence of impact and has reaffirmed the existing government support for the program. However, the findings were not available at the time of the decision to move to the second phase, making it unclear whether the impact evaluation per se merely validated an ongoing commitment or played a role in decision making.

The positive findings of the two impact evaluations the early child development program in Uganda and of the deworming for preschool children within that program were available at the time of the project's completion and cited in the ICR. However, the project, which was initially moved from a multisectoral entity to the Ministry of Health soon after it was approved, never had strong support from the latter. Further, it ran out of money and closed before being fully implemented. Child days have continued nationwide even after the end of the project, although it was unclear whether this was the result of the evaluation of the early child development program or of efforts by UNICEF. The government also introduced de-worming of preschool-age children, although it was unclear whether this decision was taken before the impact evaluation results were known.

In three countries where the evaluations found no or very small impact there was compelling evidence that the impact evaluations had no effect. An evaluation finding of small impact or no impact should not necessarily lead to the cancellation of a program—it could point to the need to introduce course corrections. However, this apparently did not occur in these three cases.

The BINP evaluation found a small positive impact of the community nutrition component on anthropometric outcomes and pointed to a number of weak links in the causal chain that could be addressed for greater impact or cost-effectiveness (White and Masset 2007). The prior evaluation sponsored by SCF pointed to some of these weak links as well, but concluded that BINP had no impact on nutrition outcomes (Hossain and others 2005). The decision to scale up the community nutrition activities in the form of the National Nutrition Program was taken at the midterm of BINP, based only on trends in project areas and before either of the impact evaluations had been issued. Nutrition has subsequently been absorbed into Bangladesh's sector-wide program.

Respondents indicated that the activities included in the community nutrition part of the program are basically unchanged and that the evaluations had had no real impact. One respondent noted, however, that at least on the part of the Bank there was much greater attention to the quality of implementation of the program, a point that was highlighted in the evaluations.

The evaluations in Bangladesh, Ecuador, and Bolivia found low impact, and the evaluations had little influence.

In Ecuador, the evaluation concluded that BDH was better targeted than its predecessor, *Bono Solidario;* the evaluation found impacts on a number of dimensions, though not specifically for HAZ (only when aggregated with two other measures). The program was targeted to all households in the two lowest quintiles of the population—40 percent of the population overall; however, the benefits were demonstrated only in the lowest quintile. The recommendation to drop the second-lowest quintile from the program was not taken, nor was the suggestion that impact might be increased by introducing conditionality based on enrollment and use of public health and education services. However, more recently, conditional transfers are being introduced in the three provinces with the highest stunting rates.[33]

The results of the impact evaluation of Bolivia's PIDI program were available in time for the ICR. The evaluation found impacts in a number of areas, though not on nutritional outcomes. Although there were political changes during the course of the project, almost from the outset it was clear that the model was extremely expensive (about $30/child/month) and not sustainable on a large scale in a country of the income level of Bolivia. As a result, the intervention initially evaluated was altered in major ways, such that what was ultimately adopted was much cheaper ($2/child/month) and sustainable, and not evaluated. The ICR noted that "all activities were ended as of December 2003 and none . . . were included or absorbed by other ongoing programs" (World Bank 2004, p. 30).

Several of the impact evaluations were reported to have increased evaluation capacity or commitment to evidence-based decision making, irrespective of the findings. These included evaluations in Colombia, Ecuador, and the Philippines. The commitment to a broader agenda of impact evaluations of social sector programs is being pursued in Colombia with World Bank support; since 2002 the number of evaluations launched by the government has risen from 3 to 30 to 46.

In Ecuador, respondents underscored that the experience with the impact evaluation greatly increased the capacity of the social sector ministry secretariat through their involvement in the design, piloting, and sample-selection phases. It reportedly led to a large change in the capacity to think about and offer impact evaluations and, although the Bank's support for this program and others was discontinued, the secretariat has reportedly launched impact evaluations on its own. In the Philippines, the evaluation—which had strong local ownership—was reported by one respondent to have had broad impacts on the design of future government programs.

Photo by Simone D. McCourtie, courtesy of the World Bank Photo Library.

On the basis of the experience with evaluating *Familias en Acción,* Colombia adopted a large program of impact evaluations for other social programs.

The scaling up of programs was often cited as evidence of the impact of the evaluations, but the features of programs that were scaled up were often substantially different from those that were evaluated. For example, the findings of the evaluation in Colombia demonstrated impact in rural areas, but the scaling up was done in urban areas. The need for an urban pilot was recognized, but in the face of an election, the intervention was expanded and the evaluation of the urban pilot was canceled.

The National Nutrition Program in Bangladesh scaled up the BINP community nutrition interventions, but some NGOs in the new areas were less experienced. In Madagascar program coverage has been extended to the whole country, but the government has dropped key elements to cut costs. These substantially different interventions have not been evaluated and their effectiveness is unknown.

Lessons

The findings in this chapter underscore important lessons for both program managers and evaluators that can guide future evaluations of the impact of large-scale government programs on nutritional outcomes.

For managers:

- **Impact evaluations of interventions that are clearly beyond the means of the government to sustain are of limited relevance.** The complexity, absolute costs, and potential sustainability of finance of the intervention should play into the decision as to whether it should be evaluated.
- **Impact evaluations are often launched for the purpose of evaluating completely new programs, but they may be equally or even more useful in improving the effectiveness of ongoing programs.** The prospects for updating an existing program with broad political and institutional support may be greater than those for a totally new program that has less ownership and may be more politically contentious.
- **There are ways of obtaining reliable results, even when randomized assignment of the intervention is not feasible for political, ethical, or practical reasons.** Correctly executed experimental designs are valuable for establishing internal validity of the evaluation, but randomization is not always possible, and even when attempted, it can be derailed in implementation of large-scale programs. Quasi-experimental methods can also be used, alone or as backup to experimental evaluations, to address the issue of the counterfactual—for example, through matching techniques and analyzing the marginal impact of longer exposure to a program.

For evaluators:

- **Evaluators would be well advised to do an ex ante risk analysis in designing impact evaluations of large government programs to anticipate how the risks to implementing the evaluation can be reduced and to chart out a contingency plan in the event that risk mitigation is not successful.** Large public nutrition programs are sensitive to political changes and budget crises; these factors should be considered in the planning of impact evaluations to maximize the success of the evaluation (beyond any project-related risk analysis).
- **Nutrition impact evaluations, in their design and analysis of the data, need to take into account the sensitivity of different age groups to the interventions.** Interventions found to be ineffective for a large age range may nonetheless be important for children at certain points in their development, particularly during gestation and in the first two years of life.
- **Evaluators must thoroughly understand the interventions being evaluated and when delivery of the intervention effectively took place.** Failure to take into account the timing of implementation can mute the measured impact of the intervention.
- **Impact evaluations need to collect rich data to document the delivery of program outputs, their quality, and their intermediate outcomes to establish the plausibility of evaluation results and to point to parts of the program that work and do not work.** The nutrition impact evaluations reviewed here have generally failed to collect sufficiently rich data, including process evaluations in parallel, to help identify what parts of the program are working and to explain why some program elements are ineffective. Too often, the lack of impact is not sufficiently followed up with an understanding of how effectiveness can be improved. Any significant impact, even a small one for a subgroup, is often hailed as evidence that the program worked, without understanding how impacts can be enhanced.
- **Evaluations need to provide evidence for timely decision making, but with sufficient elapsed time for a plausible impact to have occurred.** There is clearly tension between the need to report results quickly and to ensure that the intervention has had time to work. There are benefits to disseminating early baseline and midterm results prospectively, along with process data and intermediate outcome data that can point to changes along the results chain, even when longer-term rounds of data collection are planned.
- **Nutrition impact evaluations need to invest more in documenting the targeting and cost-effectiveness of supplemental feeding for malnourished or growth-faltering children;** the food element of the community nutrition and early child development programs often accounted for half or more of the total cost of the program. Food distribution is often politically popular, but it creates many logistical problems and is demanding of implementers, who must prevent leakage. Different delivery mechanisms for feeding need to be evaluated as well (for example, observed by a health worker versus take-home rations).
- **Evaluations of interventions to improve nutrition need to assess systematically the distribution of the benefits and the complementarities with public health and other services.** Too few evaluations assessed the extent to which the poor disproportionately benefit in relation to the nonpoor, or the impact of the availability or quality of health services on the ability of the poor to act on the information they receive on better nutrition.

Chapter 4

Photo by Curt Carnemark, courtesy of the World Bank Photo Library.

Conclusions

High rates of childhood malnutrition in developing countries are raising mortality and present long-term consequences for survivors. Progress in reducing child malnutrition has been slow, and the global food and financial crises have no doubt created setbacks. In this context, the World Bank is expanding its support for nutrition and, in parallel, has launched several new impact evaluation initiatives.

This review has attempted to inform these new efforts to improve the impact of nutrition support through a two-pronged approach.

- First, IEG reviewed the recent impact evaluation research on the effectiveness of interventions and programs in improving nutrition outcomes, focusing on child anthropometrics and birthweight. Forty-six recent nutrition impact evaluations were reviewed, representing evidence from 25 developing countries and a variety of interventions, including large-scale social programs of conditional and unconditional cash transfers, community-based nutrition, integrated health services, early child development, food transfers, de-worming, and micronutrient supplementation, among others.
- Second, IEG examined in detail the experience from impact evaluations embedded in World Bank projects that sought to affect anthropometric outcomes. Twelve impact evaluations reviewed in the first part could be linked to evaluation of Bank support to eight countries. The review examined the design, implementation difficulties, findings, and impact of the impact evaluations, based on a review of project documents, the evaluation results, and interviews with Bank staff, the evaluators, and individuals from the borrowing countries.

The overarching conclusion of the review is that *context matters.* A wide range of interventions was found to have an impact on indicators related to height, weight, wasting, and birthweight. In many settings, however, similar interventions had no effect. The magnitude of program impacts was not only difficult to compare across studies but also variable.

The findings overall do not lend themselves easily to generalizations about what works and does not work in reducing malnutrition—particularly as applied in field conditions of developing countries. Some results are based on RCTs with short results chains. But when it comes to evaluation of more complex programs implemented outside of a research setting the evaluation must document a long causal chain. Many things can go wrong, both in the quality of implementation of the intervention on the supply side and in the response of households on the demand side.

This has several implications:

- **It should not be assumed that an intervention found effective in an RCT in the medical literature will have the same effects when implemented under field conditions as part of a large program with a mix of interventions and in a population for which the underlying factors affecting malnutrition may be fundamentally different.**
- **It is important for the design of both the program and the evaluation to understand the prevailing underlying causes of malnutrition in any given setting.** When there are multiple channels and several are equally important, addressing only one of them may have limited impact.
- **Impact evaluations need to collect rich data on program service delivery and demand-side behavioral outcomes to explain nutrition impacts.** Irrespective of the evaluation design, it is critically important to understand not only *whether* the outcome is different between a treatment and comparison or control group but also *why.* When an evaluation finds no significant impact of an intervention that theoretically should have an effect, it is important to find out where in the causal chain the program broke down. This involves conducting process evaluations and collecting data to document the causal chain in parallel. In particular, many interventions involve costly food supplementation, but the functioning, targeting, and impact of food supplementation are not tracked with respect to how it contributes to outcomes.

Evaluations need to look more closely at the distribution of impacts. Very few of the evaluations reviewed examined who is benefiting and who is not. Just because malnutrition is more common among the poor does not mean that they will disproportionately benefit from a nutrition program, particularly if acting on new knowledge or different incentives relies on access to education or quality services. Very few of the evaluations assessed whether the impact differed according to the availability of complementary health services. Several found, in fact, that the children of more educated mothers are benefiting the most.

A number of lessons for development practitioners and evaluators arose from the review of impact evaluations of World Bank nutrition support. Impact evaluations should be prioritized for relevant interventions that are within the capacity and budget of the country to implement and sustain. Though most evaluations are of completely new programs, there is considerable scope for improving program effectiveness through impact evaluations of enhancement of ongoing programs.[1] There are ways of obtaining reliable results, even when randomized assignment of the interventions is not feasible.

There are many challenges to implementing evaluations of large-scale programs with a long results chain; assessing the risks to the evaluation design and implementation ex ante and planning mitigation measures can help keep an evaluation on course. Nutrition impact evaluations, in their design and analysis, need to take into account the sensitivity of different age groups to the interventions. Evaluators also need to understand exactly when delivery of the intervention effectively took place. Evaluation results need to be delivered in time to provide evidence for decision making, but with sufficient elapsed time for a plausible impact to have occurred. Impact evaluations provide a rare opportunity to document both costs and effects, yet cost-effectiveness is rarely analyzed. With these factors in mind, impact evaluations of World Bank–supported programs to affect nutrition can have a far greater impact on program effectiveness.

In sum, in approaching the impact evaluation literature and the conduct of nutrition impact evaluations, we shouldn't be asking simply, "What works?" but rather, "Under what conditions does it work, for whom, what part of the intervention works, and for how much?" These are important questions that development practitioners should be asking in reviewing the literature and that evaluators should be addressing to improve the relevance and impact of nutrition impact evaluations.

Photo by Curt Carnemark, courtesy of the World Bank Photo Library.

APPENDIX A

The Impact Evaluations Reviewed

Study[a]	Country—Program Name[b]	Intervention[c]	Study duration	Sample size	Evaluation method[d]	Significant impact on: Height, HAZ, stunting	Weight, WAZ, underweight	WHZ, wasting	Birth-weight, LBW	Inter-mediate outcome[e]	Hetero-geneity[f]
Agüero, Carter, and Woolard (2007)	South Africa—Child Support Grant	CT	18 months[g]	1,606 children aged 0–36 months	PSM with dose response	Yes					
Alderman (2007)	Uganda—Early Child Development	D, G, NE	1 year	9,073 children aged 0–60 months	DID		Yes			B, D	E, I
Alderman and others (2009)	Senegal—PRN	D ,G, M, NE, P	3 years[h]	10,378 children less than 3 years of age	DID		Yes			B, I, M, U	
Alderman and others (2006)	Uganda—Early Child Development and de-worming	D, G, M, NE	4 years[i]	27,995 children aged 1–7 years	R		Yes				
Alderman, Hoogeveen, and Rossi (2006)	Tanzania—*Partage*	F		1,140 children <5 years of age surveyed between one and four times	IV	Yes	Yes				
Armecin and others (2006)	Philippines—Early Child Development	F, G, M, NE, P, T	36 months	6,693 children aged 24–84 months	DID, PSM	Yes[j]		Yes		I, M	
Attanasio and Vera-Hernández (2004)	Colombia—*Hogares Comunitarios*	DC, F, G, M	n.a.	4,147 children aged 0–6 years	IV, PSM, quantile regression	Yes	No				E
Attanasio and others (2005)	Colombia—*Familias en Acción*	CT, F, G, M, NE, T	4 years	8,919 children aged 0–6 years (0–83 months)	PSM, DID	Yes			Yes	D, I, U	R
Barber and Gertler (2008)	Mexico—*Oportunidades*	CT, F, G, M, NE, P	20 months[k]	840 women in rural areas; mean age 29 years	R, IV				Yes	U	
Behrman and Hoddinott (2005)	Mexico—*Oportunidades*	CT, F, G, M, NE, P, T	24 months	601 children aged 4–48 months	PSM, DID	Yes			-		E, O
Behrman, Cheng, and Todd (2004)	Bolivia—PIDI	DC, F, G, M	2–25+ months	Participating children aged 6–72 months: 1,198 in the first round, 2,420 in the second round, and 364 in both rounds	PSM with dose response	No	No				
Bobonis, Miguel, and Sharma (2006)	India—Pratham Delhi Preschool Health Program	D, M	5 months	2,383 children aged 24–72 months	R, DID	No	Yes	Yes		I, M	E, G, O

Christian and others (2003)	Nepal	M	29 months[h]	4,926 pregnant women and 4,130 live-born infants	R				Yes		
Das Gupta and others (2005)	India—ICDS	Various[l]	n.a.	90,000 households; ages of children ranged from 0–4 years in the 1992 survey and 0–3 years in the 1998 survey	PSM	No	No				G, R
Friis and others (2004)	Zimbabwe	M	Starting from week 22–35 week of gestation	1,669 women; birth data available for 1,106 women	R				Yes[m]		O
Galasso and Umapathi (2009)	Madagascar—SEECALINE	F, G, M, NE, P	8 years	14,000 households and 12,367 (at follow-up); children aged 0–60 months	PSM , DID	Yes	Yes			B, D, H, I, M, P, U	E, I, R, O
Galasso and Yau (2006)	Madagascar—SEECALINE	F, G, M, NE, P, S	24 months	Children aged 0–36 months in 2,697 communities	PSM		Yes				I, O
Gertler (2004)	Mexico—*Oportunidades*	CT, F, G, M, NE, P, T	24 months	1,552 children aged 12–36 months	R	Yes				I	
Gupta and others (2007)	India	M	37–77 days	200 pregnant women randomized into treatment and control groups	R				Yes	I	
Hossain and others (2005)	Bangladesh—BINP	F, G, M, NE, P	6 years	2,338 children aged 6–23 months	Matching program and nonprogram areas	No	No	No		B, P	
Iannotti and others (2008)	Peru	M	15.6 weeks	1,295 women enrolled; 1,079 neonates measured for anthropometry at birth; 546 children included for growth analysis from birth to 12 months	R	No	Yes		No	B, D, I	G

(continued on next page)

Study[a]	Country—Program Name[b]	Intervention[c]	Study duration	Sample size	Evaluation method[d]	Significant impact on: Height, HAZ, stunting	Weight, WAZ, underweight	WHZ, wasting	Birth-weight, LBW	Inter-mediate outcome[e]	Hetero-geneity[f]
Kazianga, de Walque, and Alderman (2009)	Burkina Faso—School meals and THRs	F, THR	Not clear[n]	Covers 46 new schools that were opened in 2005–06 school year; 2,208 households (having a total of 4,140 school-age [6–15 years] children) around these schools were surveyed	R, DID	No	Yes	Yes		M	
León and Younger (2007)	Ecuador—*Bono Solidario*	CT	n.a.	5,824 households and approximately 3,000 children under 5 years of age	IV	Yes	No				O
Leroy and others (2008)	Mexico—*Oportunidades*	CT, F, G, M, NE, P, T	24 months	432 children younger than 24 months	PSM, DID	Yes	Yes	Yes			I
Linnemayr and Alderman (2008)	Senegal—PRN	D, G, M, NE, P	1 year	4,296 baseline and 6,144 endline children 0–35 months of age	PSM, DID		Yes			B, I, M, U	I
Macours, Schady, and Vakis (2008)	Nicaragua—*Atención a Crisis*	CT, F, G, M, NE, P, T	1 year[h]	3,506 children aged 0–83 months	R	No	No		No	D, U	
Maluccio and Flores (2005)	Nicaragua—RPS	CT, G, M, NE, P, T	24 months	987 children (at follow-up) 5 years and younger	R, DID	Yes	Yes	No		M, U	G, I
Masanja and others (2005)	Tanzania—IMCI	n.a.	n.a.	2,006 (1999 survey) and 1,924 (2002 survey) children 24–59 months of age	Matching IMCI and non-IMCI areas	Yes	Yes	No			
Menéndez and others (2008)	Mozambique	Sulphadoxine-pyrimethamine with insecticide-treated nets		1,027 pregnant women randomized, 990 births analyzed	R				No[o]	I, M	O
Miguel and Kremer (2004)	Kenya	D and hygiene education	2 years[p]	30,000 school children 6–18 years of age	R	Yes	No			I, M	
Morris and others (2004)	Brazil—*Bolsa Alimentação*	CT, F, G, M, NE, P, T	6+ months	1,889 children 0–84 months of age	IV[q]	No	Yes[j]				
Osrin and others (2005)	Nepal	M	5 months	1,200 pregnant women (final sample 1,052)	R				Yes		

Paxson and Schady (forthcoming)	Ecuador—BDH	CT	17 months	1,479 (baseline full sample) children 3–7 years of age	R	No				U, M	I, G
Penny and others (2005)	Peru	NE	2 years[h]	377 children followed from birth to 18 months	R	Yes	Yes	No		B, D	
Quisumbing (2003)	Ethiopia—Food Aid	FFW, FD	n.a.	1,500 households and 2,968 children 0–9 years of age	MLE	Yes[j]		Yes			I,G
Ramakrishnan and others (2003)	Mexico	M		873 pregnant women	R				No		
Rivera and others (2004)	Mexico—*Oportunidades*	CT, F, G, M, NE, P, T	24 months	650 children 0–12 months of age	R	Yes					I
Ruel and others (2008)	Haiti—World Vision Programs	F, G, M, NE	9–18 months	1,500 children per survey, 12–41 months of age	R	Yes	Yes	Yes			
Santos and others (2001)	Brazil—IMCI	NE	6 months	33 doctors with 12–13 patients; about 424 children <18 months old at baseline	R	No	Yes	Yes		B, D, M	
Schipani and others (2002)	Thailand	Mixed gardening	12 months[r]	60 households; 85 children aged 1–7 years	Matching[s]	No	No	No		M	
Schroeder and others (2002)	Vietnam—CENP	D, G, F, NE	12 months	238 children, 5–30 months of age	R	No	No	No			
Stifel and Alderman (2006)	Peru—*Vaso de Leche*	FT	n.a.	19,053 children aged 0–59 months	IV	No					
Waters and others (2006)	Peru	NE	18 months	338 children 0–18 months of age	R	Yes	No			U	
White and Masset (2007)/IEG (2005)	Bangladesh—BINP	F, G, M, NE, P		Data from three sources;[t] children 6–23 months of age	PSM	Yes	Yes	Yes		P	E, O
Yamano, Alderman, and Christiansen (2005)	Ethiopia—Food Aid	FFW, FD	n.a.	2,089 children aged 6–60 months	IV	Yes					
Zeng and others (2008)	China	M	6 months[u]	5,828 pregnant women and 4,697 live births	R				Yes	I, M	

(continued on next page)

APPENDIX A *(continued)*

Source: IEG analysis.

Notes: Blank space indicates that no data were collected. HAZ = height-for-age z-score; LBW = low birthweight; WAZ = weight-for-age z-score; WHZ = weight-for-height z-score; n.a. = not available or not mentioned in the reviewed material.

a. Studies in bold were carried out by, or in collaboration with, World Bank researchers.
b. BDH = *Bono de Desarrollo Humano;* BINP = Bangladesh Integrated Nutrition Project; CENP = Community Empowerment and Nutrition Project; ICDS = Integrated Child Development Services; IMCI = Integrated Management of Childhood Illness; PIDI = *Proyecto Integral de Desarrollo Infantil;* PRN = *Programme de Renforcement de la Nutrition;* RPS = *Red de Protección Social;* SEECALINE = *Projet de Surveillance et Éducation des Écoles et des Communautés en Matière d'Alimentation et de Nutrition Élargi*.
c. CT = cash transfer; DC = day care; D = de-worming; F = feeding; FT = food transfer; FD = free food distribution; FFW = food for work; G = growth monitoring/promotion; M = micronutrient supplement; NE = nutrition education; P = prenatal services; T = treatment of illness; THR = take-home rations.
d. DID = difference-in-difference/double difference; FE = fixed effects; IV = instrumental variables/two-stage least squares; PSM = propensity score matching, R = randomized.
e. B = breastfeeding, colostrums knowledge, practice; D = dietary intake; H = hygiene; I = illness; M = micronutrient intake or status; P = pregnancy knowledge/practice; U = health care use.
f. E = mother's education; G = child's gender; I = household income or socioeconomic status; R = region/location or place of residence; O = other.
g. The treated children in their three-year window. Values are computed based on the percentage of exposure in the three-year window reported in the study.
h. Study period.
i. Five rounds of health days in four years.
j. Negative impact on some specifications.
k. April 1998–November 1999 for the treatment group. The control group started receiving benefits in December 1999.
l. Child growth monitoring, supplementary feeding, preschool education to young children, and some basic health services to young children, pregnant women, and lactating mothers.
m. Significant impact on birthweight but not on incidence of LBW.
n. One school year.
o. No significant impact on birth outcomes. However, when the results are disaggregated by gravidity, low birthweight is lower for women with four or more pregnancies.
p. 1998 and 1999 school years. The program is phased in three groups; Group 3 (the control) received treatment in 2001.
q. The control is the group that was excluded because of "random administrative error."
r. The study considered children from mixed-gardening households as the treatment group and children from non-mixed-gardening households as the controls group. The data were collected in three different seasons of the year to account for seasonal variation.
s. Mixed-gardening and nongardening households were randomly selected and matched for comparison.
t. BINP data from Karim and others (2003) on pregnant mothers and infants (6–24 months), the Save the Children data on 1,450 children aged 6–59 months, and the Helen Keller International Nutritional Surveillance Project data of about 10,000 rural households with at least one child under 59 months prior to 2000.
u. Approximation based on the mean enrollment of gestation week, which is about 14 weeks.

APPENDIX B

Impact Evaluations of Height, Height for Age, and Stunting

Study	Country—Program[a]	Intervention[b]	Evaluation method[c]	Baseline level	Impacts[d]	Heterogeneity of impacts: Income/Poverty level	Heterogeneity of impacts: Maternal education	Heterogeneity of impacts: Other	Remarks
Agüero, Carter, and Woolard (2007)	South Africa—Child Support Grant	CT	PSM with dose response	n.a.	*Children's age: 0–36 months* Significant impact on HAZ, and the gains are maximum when treatment covers around two-thirds of the nutritional window or the first three years of the child's life. Treatment effect is analyzed by age group: <12 months, 12–24 months, and 24–36 months. The effect is the largest when treatment is started at the youngest age.				
Alderman, Hoogeveen, and Rossi (2006)	Tanzania	F	IV	n.a.	*Children's age: 0–60 months* Availability of feeding post in the village is significant in the HAZ equation.				
Armecin and others (2006)	Philippines—ECD	D, DC, F, G, M, NE, P, T	PSM, DID	HAZ T = –1.57 C = –1.67 Stunting (%) T = 34.9 C = 38.7	*Children's age: 24–84 months* HAZ T = –1.79 C = –1.88 No impact in 14 of the 15 specifications. One negative impact. Stunting (%) T = 43.4 C = 48.6 Mixed impact on stunting. Of the total 15 specifications, 3 show improvement, 5 show worsening, and 7 show no impact.				The study estimates 15 models for each outcome, that is, five age groups (2, 3, 4, 5, and 6+ years) by three duration-of-exposure categories (4–12 months, 13–16 months, 17+ months).

(continued on next page)

APPENDIX B *(continued)*

Study	Country—Program[a]	Intervention[b]	Evaluation method[c]	Baseline level	Impacts[d]	Heterogeneity of impacts: Income/Poverty level	Heterogeneity of impacts: Maternal education	Heterogeneity of impacts: Other	Remarks
Attanasio and others (2005)	Colombia—Familias en Acción	CT, F, G, M, NE, P, T	PSM, DID	n.a.	*Children's age: 0–83 months* Program impact on HAZ by age group: <24 months = 0.161* 24–48 months = 0.011 (ns) >48 months = 0.012 (ns)				
Attanasio and Vera-Hernández (2004)	Colombia—*Hogares Comunitarios*	DC, F, G, M	IV, PSM, quintile regression	HAZ for children <6 years = −1.24	*Children's age: 0–72 months* Measured impacts on HAZ for three measures of program take-up: Attendance = 0.486** No. of months = 0.013** Exposure = 0.78**		Children whose mothers had no education benefited most.		Identification based on IV; attendance is number of months in the program; exposure is number of months adjusted by age of the child
Behrman, Cheng, and Todd (2004)	Bolivia—PIDI	DC, F, G, M	PSM with dose response	n.a.	*Children's age: 6–72 months* Insignificant impact on height for all age groups. Cumulative impact estimated for combined age cohorts (6–24, 25–36, 37–41, 42–58, 59+ months) and duration of exposure (1–6, 7–12, 13–18, 19–24, 25+ months).				
Behrman and Hoddinott (2005)	Mexico—*Oportunidades*	CT, F, G, M, NE, P, T	R, FE	HAZ[e] 4–12 months = −0.949 12–36 months = −1.928 36–48 months = −1.947	*Children's age: 4–48 months* Height in cm 4–12 months = 0.503 (n.s.) 12–36 months = 1.016** 36–48 months = −0.349 (n.s.)		Significant program impact for children 12–36 months whose mothers have >5 years of schooling.	Other significant impacts of household head characteristics, not by community infrastructure.	
Bobonis, Miguel, and Sharma (2006)	India—Pratham Preschools Delhi	D, M	R, DID	HAZ T = −0.79 C = −0.45. The treatment group had statistically significantly lower z-scores at baseline.	*Children's age: 24–72 months* HAZ declined by 0.19 (n.s.)				Treatment and control groups were not balanced at baseline. Randomization did not eliminate the HAZ differences between them. Estimated DID to control for initial differences.

Das Gupta and others (2005)	India—ICDS	F, G, M, NE, P, T	PSM	n.a.	*Children age 0–4 years in the 1992 survey, and 0–3 years in the 1998 survey* HAZ 1992: T = –1.877, C = –1.933 1998 : T = –1.807, C = –1.832 No impact.			Impact for boys' sample in 1992, not in 1998. For girls' sample no impact in either year. No differential impacts by region.	All results reported here are based on matched treated and control groups. The study is not longitudinal; based on two cross-sectional studies from 1992 and 1998.
Gertler (2004)	Mexico—*Oportunidades*	CT , F, G, M, NE, P, T	R	n.a.	*Children's age: 12–36 months* Children in the treatment group grew about 0.96 cm more than those in the control group.*** Treatment children were 8.6% less likely to be stunted (n.s.)				
Galasso and Umapathi (2009)	Madagascar—SEECALINE	F, G, M, NE, P	PSM	HAZ (1997–98) T = –1.942 C = –1.831	*Children's age: 0–60 months* HAZ (2004) T = –2.017 C = –1.943 HAZ (2007) T = –2.012 C = –1.903 Deteriorated in both program and nonprogram areas; deterioration significantly less in program areas.	HAZ by poverty incidence: Lowest poverty incidence = 0.287*** Middle = –0.001 (n.s.) Highest poverty incidence = 0.020 (n.s.) Stunting by poverty incidence: Lowest poverty incidence = –0.084*** Middle = 0.022 (n.s.) Highest poverty incidence = –0.032 (n.s.)	2004 HAZ: Unschooled = 0.044 (n.s.) Primary = 0.097 (n.s.) Secondary & higher = 0.141* 2007 HAZ: Unschooled = –0.151 (n.s.) Primary = 0.120 (n.s.) Secondary & higher = 0.323** 2004 Stunting: Unschooled = –0.005 (n.s.) Primary = –0.019 (n.s.) Secondary & higher = –0.051 (n.s.) 2007 Stunting: Unschooled = 0.104 (n.s.) Primary = 0.005 (n.s.) Secondary & higher = –0.103 (n.s.)	HAZ Proximity to a rural road: Yes = 0.118* No = 0.061 (n.s.) Hospital Yes = 0.160* No = 0.084 (n.s.) No differential impact in communities with and without electricity or access to a safe water source.	Deterioration concentrated among children with unschooled mothers.

(continued on next page)

Study	Country—Program[a]	Intervention[b]	Evaluation method[c]	Baseline level	Impacts[d]	Heterogeneity of impacts: Income/Poverty level	Heterogeneity of impacts: Maternal education	Heterogeneity of impacts: Other	Remarks
Hossain and others (2005)	Bangladesh—BINP	F, G, M, NE, P	Matching program and nonprogram areas	n.a.	*Children's age: 6–23 months* Severe stunting (<–3 SD) (%) T = 11.6 C = 12.4 Moderate Stunting: (between –2 SD and –3 SD) (%) T= 27.5 C = 27.6 No significant program impact found.				
Iannotti and others (2008)	Peru	M	R		*Children's age: 0–12 months* No impact on linear growth.				Intervention during pregnancy. Treatment = zinc + iron + folic acid. Control = iron + folic acid. Neonates followed from birth to 12 months for growth analysis.
Kazianga, de Walque, and Alderman (2009)	Burkina Faso	F	R, DID	HAZ all children School meals = –2.351 THRs = –2.086 Control = –2.317 Stunting (% of all children) School meals = 50.9 THRs = 60.0 Control = 61.7	*Children's age: 6–60 months* Impact on HAZ *School meals* 6–60 months = –0.19 (n.s.) 12–60 months = –0.135 (n.s.) *Take-home rations* 6–60 months = 0.212 (n.s.) 12–60 months = –0.189 (n.s.)				

Leon and Younger (2007)	Ecuador—*Bono Solidario*	CT	IV	n.a.	*Children's age: 0–60 months* The Bono increases expenditure by about 11%. Observed mean value for HAZ is –1.081. When Bono income is set to zero, mean value for HAZ is –1.121. In other words, the grant improves HAZ by 0.04. Statistically significant but small.				The impact of Bono income works through the expenditure per capita and the income share Using two-stage least squares (2SLS) estimates, a doubling of household expenditures per capita would increase HAZ by 0.85. Targeting the Bono cash grant to mothers was no more efficacious at reducing malnutrition than was other household income.
Leroy and others (2008)	Mexico—*Oportunidades*	CT, F, G, M, NE, P, T	PSM, DID	Height (cm) T = 70.9 C = 70.2 HAZ T =–1.9 C = –1.4	*Children's age: 0–24 months* Height gain (cm) 0.47(n.s.) HAZ gain 0.11(n.s.) Height gain (cm) 0–6 months = 1.53 ** 6–12 months = 0.73 (n.s.) 12–24 months = –0.07 (n.s.) HAZ gain 0–6 months = 0.41** 6–12 months = 0.23 (n.s.) 12–24 months = 0.02 (n.s.)	Height gain (cm) T1 = 0.86(n.s.) T2 = 0.22 (n.s.) T3 = 0.74 (n.s.) HAZ gain T1 = 0.27** T2 = 0.00 (n.s.) T3 = 0.13 (n.s.)			Income/poverty levels given in tercile (T1 to T3); T1 is the poorest.
Macours, Schady, and Vakis (2008)	Nicaragua—*Atención a Crisis*	CT , F, G, M, NE, P, T	R	HAZ Children 0–5 months T = –1.28 C = –1.10	*Children's age: 12–35 and 60–83 months* HAZ decreased by 0.052 (n.s.) Treatment effect 12–35 months = –0.116 (n.s.) 60–83 months = –0.013 (n.s.)				Impacts on other age groups calculated by Fiszbein and Schady (2009), using Macours data: 0–23 months = –0.14 (n.s.) 24–47 months = –2 (n.s.) 48–71 months = –0.03 (n.s.)

(continued on next page)

APPENDIX B *(continued)*

Study	Country—Program[a]	Intervention[b]	Evaluation method[c]	Baseline level	Impacts[d]	Heterogeneity of impacts			Remarks
						Income/ Poverty level	Maternal education	Other	
Maluccio and Flores (2005)	Nicaragua—RPS	CT , G, M, NE, P, T	R, DID	Stunting (%) T = 39.8 C = 39.5 HAZ T = –1.73 C = –1.77	*Children's age: 0–60 months* Stunting* T = 36.5 C = 41.7 HAZ–no impact T = –1.63 C = –1.8	Significant effect of the program on HAZ among the poor: 0.22.		No diferential effects by gender.	
Masanja and others (2005)	Tanzania—IMCI	Various	Compared trends and inequalities of health using two surveys	Stunting (%) T = 59 C = 51	*Children's age: 24–59 months* Stunting (%)** T = 43 C = 40				Health equity improved.
Miguel and Kremer (2004)	Kenya	D, Hygiene	R	n.a.	*Children's age: 72–216 months* HAZ T = –1.13 C = 1.22 Mean difference = 0.09*				Authors argue that the impact is the result of the use of antihelminths, not of the health education component of the program because there is no significant difference in behavior.
Morris and others (2004)	Brazil—*Bolsa Alimentação*	CT, F, G, NE, P, M, T	R	n.a.	*Children's age: 0–84 months* Program impact for all children younger than 84 months was –0.11 (a worsening outcome) but not significant. Impact on HAZ of children aged <24 months = –0.11 (n.s.) 24–48 months = –0.19 (n.s.) 48-84 months = –0.04 (n.s.)				

Paxson and Schady (2009)	Ecuador—BDH	CT, F, G, M, NE, P, T	R, DID	HAZ T = –1.22 C = –1.20	*Children's age: 36–84 months* No impact on height or HAZ.	On average, children in the lowest expenditure quartile have 24.3% higher physical outcomes than those in the control group.		Program effects on physical measures are consistently larger among girls than boys both for poorest and relatively better-off children. In some cases the gender differences are significant.	Baseline levels and impacts on HAZ by age are taken from Fiszbein and Schady (2009). Physical outcomes include three measures of physical development: the child's hemoglobin level, height for age, and fine motor control
Penny and others (2005)	Peru	NE	R	n.a	*Children's age: 0–18 months* Height at age 18 months (cm) T = 79.36 C = 78.29 Adjusted difference = 0.714 (*p* = 0.014) HAZ at 18 months T = –0.81 C = –1.19 Adjusted difference = 0.272 (*p* = 0.002)				Birth cohort followed to 18 months
Quisumbing (2003)	Ethiopia—Food Aid	FD, FFW	Heckman maximum likelihood estimate	n.a.	*Children's age: under 9 years* Children <5 in low-asset households: FFW has direct negative impact on HAZ. FD and all (FD + FFW) do not have impact on HAZ Children 5–8 in low-asset households: Food aid does not have an impact on HAZ; The direct and lagged effects of all FFW, FW, and FFW + FD are insignificant.	Impact reported by low and high asset households. (See the column to the left)		1. Gender of the child. Children <5, *In low-asset house-holds,* FFW and FD have no dif-ferential impact. The total value (FD + FFW) has negative lagged impact on girls. *In high-asset households*, no differential im-pact of food aid for all children.	Main message is that HAZ is not responsive to food aid in the short run.

(continued on next page)

APPENDIX B *(continued)*

Study	Country—Program[a]	Intervention[b]	Evaluation method[c]	Baseline level	Impacts[d]	Heterogeneity of impacts			Remarks
						Income/ Poverty level	Maternal education	Other	
Quisumbing *(Continued)*					Children <5 in high-asset households: Food aid has no impact on HAZ; The direct and lagged effects of all FFW, FD, and FFW + FD are insignificant. Children 5–8 in high-asset households: FFW has no impact. FD as well as total value (FD + FFW) have lagged but negative impact on HAZ.			2. Gender of the recipient In *low-asset households*, gender of the child interacted with the gender of the recipient (GirlsXFemale recipient) not significant for 5–8 year olds and inconsistent (positive and negative impact) for younger children. In *high-asset households*, the gender of the recipient has no effect in either group of children.	
Rivera and others (2004)	Mexico—*Oportunidades*	CT, F, G, M, NE, P, T	R	HAZ 0–6 months: T = −0.45 C = −0.25 6–12 months: T = −1.04 C = −1.03	*Children's age: 0–12 months* Significant impact on height for some children. Children (0–6 months) in the treatment group grew about 1.1 cm more than those in the control group (26.4 cm in the intervention group vs. 25.3 cm in the control group).	Impact larger for children from the poorest households			
Ruel and others (2008)	Haiti (two World Vision programs on maternal and child health and nutrition)	F, G	R	HAZ Preventive = −1.69 Recuperative = −1.65 Stunting (%) Recuperative = 37.4 Preventive = 36.7	*Children's ages: 12–41 months* HAZ Preventive = −1.53 Recuperative = −1.67 By age group, significant (higher for preventive group) for age group 24–35 months, but not for 12–23 months or 36–41 months. Stunting Significantly lower (by 4 percentage points) in the preventive communities.				

Santos and others (2001)	Brazil	NE	R	HAZ All: T = –0.23 C = –0.04	*Children's age: 0–18 months* Impact on length (cm) All: –0.12 (n.s.) *By age group* 0–6 months = –0.27 (n.s.) 6–12 months = –0.46 (n.s.) 12–18 months = 0.38 (n.s.) Impact on HAZ All: –0.04 (n.s.) *By age group* 0–6 months = –0.01 (n.s.) 6–12 months = –0.13 (n.s.) 12–18 months = –0.17 (n.s.)				
Schipani and others (2002)	Thailand	Mixed gardening	Matching program and nonprogram areas	n.a.	*Children's age: 12–84 months* No impact Rainy season (%): T = 20 C = 26.6 Cool season (%): T = 16.7 C = 23 Hot season (%): T = 22 C = 34			Results reported by season. Not significant in any season.	No baseline information is used in the analysis.
Schroeder and others (2002)	Vietnam	D, F, G, NE	R, multivariate	HAZ T = –1.65 C = –1.67 Stunting (%) T = 35.5 C = 42.9	*Children's age: 5–30 months* HAZ T = –1.66 C = –1.66 (Difference n.s.) Stunting (%) T = 36.0 C = 33.1 (Difference n.s.)				Children both in the treatment and control groups were de-wormed.
Stifel and Alderman (2006)	Peru—*Vaso de Leche*	FT	IV	n.a.	*Children's age: 0–60 months* Negative but insignificant impact in all models. The parameter estimates are also small in magnitude.				
Waters and others (2006)	Peru	NE	Multivariate	HAZ T = –0.50 C = –0.50 Stunting T = 0.05 C = 0.04	*Children's age: 0–8 months* HAZ The coefficient of intervention compared with control is 0.31 and significant. Stunting The coefficient of intervention compared with control is 0.33 and significant. The intervention prevented 11.1 cases of stunting per 100 children.				Significant differences between treatment and control in 2 of the 11 baseline characteristics. The treatment group was better in terms of proportion of mothers completing secondary education and economic quintile.

(continued on next page)

APPENDIX B *(continued)*

Study	Country—Program[a]	Intervention[b]	Evaluation method[c]	Baseline level	Impacts[d]	Heterogeneity of impacts: Income/Poverty level	Heterogeneity of impacts: Maternal education	Heterogeneity of impacts: Other	Remarks
White and Masset (2007)/IEG (2005)	Bangladesh—BINP	F, G, NE, P, M	PSM	n.a.	*Children's age: 6–23 months* HAZ Midterm: 0.10*** Endline: 0.08**				Average treatment effect on the treated, one-to-one matching (IEG 2005, p. 170) Endline HAZ n.s. in any other models.
Yamano, Alderman, and Christiansen (2005)	Ethiopia—Food Aid	FD, FFW	IV	n.a.	*Children's age: 0–60 months* Food aid has a positive impact on height of children 6–24 months. Impact more than doubles when program placement effects are controlled using IV (coefficient: 0.42 vs. 0.09). Equivalent to 2 cm faster growth in 6-month period among children in food aid communities. No impact on older groups (25–60 months).				

Source: IEG analysis.

Notes: C = control group; DID = difference-in-difference; HAZ height-for-age z-score; T = treatment group. Blank space indicates that no data were collected; n.a. = not available or not mentioned in the reviewed material; n.s. = not statistically significant. Statistical significance: * = $p \leq 0.10$, ** = $p \leq 0.05$, *** = $p \leq 0.01$.

a. BDH = *Bono de Desarrollo Humano;* BINP = *Bangladesh Integrated Nutrition Project;* ECD = early childhood development program; ICDS = Integrated Child Development Services; IMCI = Integrated Management of Childhood Illness; PIDI = *Proyecto Integral de Desarrollo Infantil;* RPS = *Red de Protección Social;* SEECALINE = *Projet de Surveillance et Éducation des Écoles et des Communautés en Matière d'Alimentation et de Nutrition Élargi.*

b. CT = cash transfer; DC= day care; D = de-worming; F = feeding; FD = free food distribution; FFW = food for work; FT = food transfer; G = growth monitoring; M = micronutrient supplement; NE = nutrition education; P = prenatal services; T = treatment of illness; THR = take-home rations.

c. DID = difference-in-difference/double difference; FE = fixed effects; IV = instrumental variable/2SLS; PSM = propensity score matching, R = randomized.

d. cm = centimeter; SD = standard deviation.

e. Computed based on HAZ information reported in the study.

APPENDIX C

Impact Evaluations of Weight, Weight for Age, and Underweight

Study	Country—Program[a]	Intervention[b]	Evaluation Method[c]	Baseline	Impacts[d]	Heterogeneity of impacts: Income/poverty level	Heterogeneity of impacts: Maternal education	Heterogeneity of impacts: Other	Remarks
Alderman (2007)	Uganda—ECD	D, G, M, NE, P	DID	WAZ T = −1.052 C = −1.060	*Children's age: 0–60 months* WAZ T = −1.105 C = −1.108 On average, there was no program impact. However, WAZ significantly improved for children <12 months in the T compared with the C (p = 0.08).	No differential impacts by household expenditure.	No differential impacts.		
Alderman, Hoogeveen, and Rossi (2006)	Tanzania—Partage	F	IV	n.a.	*Children's age: 0–60 months* Availability of feeding post in the village is significant in the WAZ equation.				
Alderman and others (2006)	Uganda ECD/De-worming	D, G, M, NE, P	R		*Children's age: 12–84 months* An increase in weight gain of about 10% (about 166 g per child per year, 95%, CI 16 to 316 g) above expected weight gain when treatments were given twice a year, and an increase of 5% when the treatment was given annually.				
Alderman and others (2009)	Senegal—PRN	D, G, M, NE, P	DID		*Children's age: 0–36 months* Children in program villages are less likely to be underweight (odds ratio 0.83, 95% CI: 0.686, 1.000).				
Attanasio and Vera-Hernández (2004)	Colombia—*Hogares Comunitarios*	DC, G, F, M	IV, quintile regression	WAZ = −0.80 Results pooled from baseline (2002) and follow-up (2003).	*Child age: 0–72 months* Impact computed for three measures of program take-up: attendance (participation), number of months in the program, and exposure (number of months adjusted by age of the child). Attendance = 0.247 (n.s.) No. of months = 0.001 (n.s.) Exposure = 0.132 (n.s.)				

(continued on next page)

APPENDIX C *(continued)*

Study	Country—Program[a]	Intervention[b]	Evaluation Method[c]	Baseline	Impacts[d]	Heterogeneity of impacts			Remarks
						Income/ poverty level	Maternal education	Other	
Behrman, Cheng, and Todd (2004)	Bolivia—PIDI	DC, G, F, M	PSM with dose response	n.a.	*Children's age: 6–72 months* Insignificant impact on weight for all age groups (6–24, 25–36, 37–41, 42–58, 59+ months) and durations of exposure (2–6, 7–12, 13–18, 19–24, 25+ months).				
Bobonis, Miguel, and Sharma (2006)	India—Pratham Delhi)	D, M	R, DID	WAZ T = –1.41 C = –1.15 The difference was statistically significant.	*Children's age: 24–72 months* WAZ improved by 0.31. Child weight gains—roughly 0.5 kg (1.1 pounds) on average—in the treatment schools relative to comparison schools.				T and C groups were not balanced at baseline. Randomization did not eliminate the WAZ differences between T and C. Estimated DID to control for initial differences.
Das Gupta and others (2005)	India—ICDS	Various	PSM	n.a.	*Children's ages: 0–4 years in the 1992 survey, and 0–3 in the 1998 survey* WAZ 1992 T = –1.917 C = –1.873 1998 T = –1.789 C = –1.788 No impact.			No differential impact by gender. Some impacts by region, but negative.	Baseline not available. Results reported here are based on matched treatment and control groups. Impacts seen at two different points in time, in 1992 and 1998.

Galasso and Umapathi (2009)	Madagascar—SEECALINE	F, G, M, NE, P	PSM	WAZ T = –1.789 C = –1.667	*Children's age: 0–60 months* WAZ T = –1.584 C = –1.574 The program improved WAZ by 0.15–0.22, and the incidence of underweight declined by 5.2–7.5 percentage points.	WAZ: Impact by incidence of poverty: Lowest poverty incidence = 0.330*** Middle = 0.001 (ns); Highest poverty incidence = 0.144** Underweight: Lowest poverty incidence = –0.130*** Middle = 0.001 (n.s.) Highest poverty incidence = –0.041 (n.s.)	2004 WAZ: Unschooled = 0.136 (n.s.) Primary = 0.121** Secondary &higher = 0.214*** 2004 Underweight: Unschooled = –0.04 (n.s.) Primary = –0.03 (n.s.) Secondary & higher = –0.111*** 2007 WAZ: Unschooled = –0.05 (n.s.) Primary = 0.152* Secondary & higher = 0.279** 2007 Underweight Unschooled = 0.008 (n.s.) Primary = –0.044 (n.s.) Secondary & higher = –0.07 (n.s.)	WAZ and underweight: Impacts are also greater with respect to proximity to rural road, hospital, electricity, and type of water source in the community.	Overall, differential impacts by maternal education and community socioeconomic characteristics. Positive trends for WAZ are observed only for educated mothers. Impacts widened when socioeconomic disparities are reinforced. Nutritional gains are larger in better-off communities.
Galasso and Yau (2006)	Madagascar—SEECALINE	F, G, M, NE, P	PSM	Underweight 0–6 months = 0.31 7–12 months = 0.46 13–36 months = 0.51	*Children's age: 0–36 months* Results are reported by duration of exposure and age. Statistically significant reductions in both cases. Reduction of malnutrition by child age: For one-year exposure (%) 0–6 months = 8 7–12 months = –4 13–36 months = 0 For two-year exposure (%) 0–6 months = 9 7–12 months = 8 13–36 months = 8	"Sites with a higher poverty rate have higher returns to exposure over two years relative to sites with lower poverty rates." (p. 22)		Program effect responsive to duration of exposure.	On average, there are positive returns to exposure.

(continued on next page)

Study	Country—Program[a]	Intervention[b]	Evaluation Method[c]	Baseline	Impacts[d]	Heterogeneity of impacts: Income/poverty level	Heterogeneity of impacts: Maternal education	Heterogeneity of impacts: Other	Remarks
Hossain and others (2005)	Bangladesh—BINP	F, G, M, NE, P	Matching program and nonprogram areas	n.a.	Severely low WAZ (<–3 z-scores) (%) T = 11.4 C = 12.1 Moderately low WAZ (≥–3 and <–2 z-scores) (%) T = 35.2 C = 36.3				T = project area C = nonproject area
Iannotti and others (2008)	Peru	M	R		*Children's age: 0–12 months* Positive impact on weight. Treatment group was heavier by 0.58 kg.***			Gender affects impact, but result not reported.	Compared zinc + iron + folic acid versus iron + folic acid in pregnant women. Neonates followed from birth to 12 months for growth analysis.
Kazianga, de Walque, and Alderman (2009)	Burkina Faso	F	R	WAZ all children School meals = –2.202 Take-home rations = –2.521 Control = –2.394 Underweight (% of children) All children School meals = 52.6 Take-home rations = 56.2 Control = 55.3	*Children's ages: 6–60 months and 6–10 years* Impact on WAZ *School meals* 6–60 months = 0.219** 12–60 months = 0.172 6–10 years = 0225* *Take-home rations* 6–60 months = 0.355* 12–60 months = 0.376* 6–10 years = 0.153 (n.s.)				
Leon and Younger (2007)	Ecuador—*Bono Solidario*	CT	IV	n.a.	*Children's age: 0–60 months* There is no impact of the Bono share of household expenditure on WAZ. Household expenditure per capita does have a significant impact, however, and Bono income increases expenditure by 11%.				The impact of Bono income works through the expenditure-per-capita variable and the income-share variable.

Leroy and others (2008)	Mexico—*Oportunidades*	CT, F, G, M, NE, P, T	PSM, DID	Weight (kg) T = 8.62 C = 8.46	*Children's age: 0–24 months* Weight gain 0.202 kg (ns) Weight gain by age (kg) 0–6 months = 0.763 (n.s.) 6–12 months = 0.026 (n.s.) 12–24 months = 0.09 (n.s.)	Weight gain (in terciles [T]) T1 = 0.523** T2 = 0.079 (n.s.) T3 = 0.138 (n.s.)			Income/poverty levels given in tercile (T1–T3); T1 is the poorest.
Linnemayr and Alderman (2008)	Senegal—PRN	D, G, M, NE, P	DID	WAZ score T = −1.352 C = −1.276	*Children's age: 0–36 months* No significant program impact on WAZ overall. A dummy variable for full exposure is created for younger children (0–6 months). Impact is significant for this group.	No impact by wealth.			
Macours, Schady, and Vakis (2008)	Nicaragua—*Atención a Crisis*	CT, G, M, NE, P, T	R	Average WAZ for children 0–5 months old at baseline were −0.91 and −1.06 for C and T groups, respectively. The difference is not statistically significant (p = 0.168).	*Children's age: 12–35 months and 60–83 months* WAZ decreased by 0.052 (n.s.) Impact on WAZ by age 12–35 months = −0.114 (n.s.) 60–83 months = 0.046 (n.s.)				
Maluccio and Flores (2005)	Nicaragua—RPS	CT, F, G, M, NE, P, T	R, DID	Baseline in 2000. Proportion of children under 5 who were underweight was 13.7% in the treatment group and 14.3% in the control group.	*Children's age: 0–60 months* The program also improved underweight in the intervention areas. Percentage underweight declined from 13.7 to 9.8 in the treatment group and increased from 14.3 to 16.6 in the control group.**			No differential effects by gender.	
Masanja and others (2005)	Tanzania—IMCI	Various	Compared trends and inequalities of health for children 24–59 months using two surveys	Underweight (%) In 1999 T = 30 C = 27	*Children's age: 24–59 months* Underweight (%)** in 2002 T = 23 C = 19 Difference in differences (IMCI districts [T] minus non-IMCI districts [C]) = 0.044.				The program started earlier and the authors compared trends and inequalities of health using surveys conducted in 1999 (n = 2,006) and 2002 (n = 1,924) for children under 5 years of age.

(continued on next page)

Study	Country—Program[a]	Intervention[b]	Evaluation Method[c]	Baseline	Impacts[d]	Heterogeneity of impacts			Remarks
						Income/ poverty level	Maternal education	Other	
Miguel and Kremer (2004)	Kenya	D, hygiene	R	WAZ T = –1.39 C = –1.40	*Children's age: 72–216 months* WAZ score T = –1.25 C = –1.25 (no difference)				Estimated externalities by treatment status. Difference was not statistically significant.
Morris and others (2004)	Brazil—*Bolsa Alimentação*	CT, F, G, M, NE, P, T	IV		*Children's age: 0–84 months* WAZ for all children in beneficiary households was 0.13 lower. No significant impact on WAZ is found when the sample is disaggregated by age groups. For three age groups, <24 months, 24–48 months, and 48–84 months, project impacts on WAZ were 0.25, –0.11, and –0.08, respectively.				The authors indicated that negative impact might be "due to perceptions that benefits would be discontinued if the child started to grow well."
Penny and others (2005)	Peru	NE	R	n.a.	*Children's age: 0–18 months* Mean weight at 18 months (kg) T = 10.77 C = 10.48 Adjusted difference = 0.199 (p = 0.093) WAZ at 18 months T = –0.33 C = –0.62 Adjusted difference = 0.194 (p = 0.041)				Birth cohort followed from birth to age 18 months.

Ruel and others (2008)	Haiti	F, G, M, NE, P	R	WAZ Preventive = −0.97 Recuperative = −1.02	*Children's age: 12–41 months* WAZ Preventive = −0.96 Recuperative = −1.2 At the end of three years, children from preventive communities had significantly higher mean WAZ (+0 • 24) than the recuperative group. By age group, the difference was significant (higher for preventive group) for age groups 12–23 months and 24–35 months, but not for 36–41 months. Underweight Significantly lower (6 percentage points) in the preventive model.				
Santos and others (2001)	Brazil	NE	R	WAZ All: T = 0.06 C = 0.31	*Children's age: 0–18 months* Impact on Weight (kg) All: −0.01 (n.s.) By age group 0–6 months = −0.16 (n.s.) 6–12 months = 0.07 (n.s.) 12–18 months = 0.34** Impact on WAZ All: 0.07 (n.s.) By age group 0–6 months = −0.09 (n.s.) 6–12 months = 0.13 (n.s.) 12–18 months = 0.31**				
Schipani and others (2002)	Thailand	Mixed gardening	Matching program and nonprogram areas	n.a.	*Children's age: 1–84 months* No significant impact on underweight. No impact on wasting. Results for % of WAZ ≤2 in three seasons: Rainy season T = 10.8 C = 16.6 Cool season T = 6.8 C = 16.6 Hot season T = 14.8 C = 20.6				No baseline information is used in the analysis. The study is based on single difference from a cross-sectional data collected in three seasons.

(continued on next page)

APPENDIX C *(continued)*

Study	Country—Program[a]	Intervention[b]	Evaluation Method[c]	Baseline	Impacts[d]	Heterogeneity of impacts: Income/poverty level	Heterogeneity of impacts: Maternal education	Heterogeneity of impacts: Other	Remarks
Schroeder and others (2002)	Vietnam—CENP	D, F, G, NE,	R, multivariate	WAZ T = −1.51 C = −1.68 Underweight (WAZ < −2 SD) T = 30.3 C = 35.3	*Children's age: 5–30 months* WAZ T = −1.92 C = −2.06. The difference is not significant ($p = 0.19$). Underweight T = 46.5 C = 55.9. The difference is not significant ($p = 0.15$). Children who were younger (<15 months) and more malnourished (less than −2 WAZ) at baseline had significantly better growth than similarly young, malnourished children in the control group.				
Waters and others (2006)	Peru	NE	R, multivariate	WAZ[e] T = −0.26 C = −0.34 Underweight T = 3 C = 4	*Children's age: 0–18 months* The coefficients of the intervention on WAZ and underweight regressions are insignificant.				
White and Masset (2007)/IEG (2005)	Bangladesh—BINP	F, G, M, NE, P	PSM	n.a.	*Children's age: 6–23 months* WAZ Midterm: 0.12*** Endline: 0.09**	Interactions with wealth quantiles at endline n.s.	Interactions with mother's education at endline n.s.		Average treatment effect on the treated, one-to-one matching (IEG 2005, p. 170). Interactions reported in IEG 2005, p. 177.

Source: IEG analysis.

Note: C = control group; T = treatment group; WAZ = weight-for-age z-score; n.a. = not applicable; n.s. = not significant. Statistical significance: * = $p \leq 0.10$; ** = $P \leq 0.05$; *** = $p \leq 0.01$.

a. BINP = Bangladesh Integrated Nutrition Project; CENP = Community Empowerment and Nutrition Project; ECD = early childhood development program; ICDS = Integrated Child Development Services; IMCI = Integrated Management of Childhood Illness; PIDI = *Proyecto Integral de Desarrollo Infantil;* PRN = *Programme de Renforcement de la Nutrition;* RPS = *Red de Protección Social;* SEECALINE = *Projet de Surveillance et Éducation des Écoles et des Communautés en Matière d'Alimentation et de Nutrition Élargi.*

b. CT = cash transfer; DC = day care; D =de-worming; F = feeding; FD = free food distribution; FFW = food for work, FT = food transfer; G = growth monitoring; M = micronutrient supplement; NE = nutrition education; P = prenatal services; T = treatment of illness.

c. DID = difference in difference/double difference; FE = fixed effects; IV = instrumental variable/2SLS; PSM = propensity score matching, R = randomized.

d. C = nonproject area; CI = confidence interval; g = gram; kg = kilogram; T = project area; WAS = weight-for-age z-score.

e. Significant differences between treatment and control in 2 of the 11 baseline characteristics included in the study. The treatment group is better off in terms of proportion of mothers completing secondary education and economic quintile.

APPENDIX D

Impact Evaluations of Weight for Height and Wasting

Study	Country—Program[a]	Intervention[b]	Evaluation Method[c]	Baseline	Findings[d]	Heterogeneity of impacts			Remarks
						Income/ poverty level	Maternal education	Other	
Armecin and others (2006)	Philippines—Comprehensive ECD	D, DC, F, G, M, NE, P, T	PSM, DID	WHZ T = –0.638 C = –0.696 Wasting (%) T = 7.4 C = 6.1	*Children's age: 24–84 months* WHZ T = –0.355 C = –0.526. Wasting (%) T = 2.4 C = 2.5. Results reported by age group and duration of exposure. Overall, mixed but predominantly positive results. By age group: 2- and 3-year-olds, mixed (negative and positive), or no impact. 4-, 5-, and 6+ year olds, positive impacts. The mean impact of all outcomes is the biggest for the youngest age group (2 years). This group has also the largest number of significant cases.				The study estimates 15 models for each outcome, that is, five age groups (2, 3, 4, 5, and 6+ years) by three duration-of-exposure categories (4–12 months, 13–16 months, 17 months). The authors summarize the number of positive and significant specifications for WHZ and wasting, fine and gross motor skills, cognitive skills, and self-help, as well as expressive and receptive language.
Bobonis, Miguel, and Sharma (2006)	India—Pratham Delhi	D, M	R, DID	WHZ T = –1.12 C = –1.02	*Children's age: 24–72 months* WHZ Improved by 0.52*** WHZ by age group 2–3 years = 0.57*** 4–6 years = 0.46**		School <3 years = 0.63** >3 years = 0.37*	Gender Boys = 0.29* Girls = 0.67*** Probability of being anemic at baseline High = 0.62** Low = 0.32*	
Hossain and others (2005)	Bangladesh—BINP	F, G, M, NE, P	Matching program and nonprogram areas	n.a.	*Children's age: 6–23 months* Severe wasting (WHZ < –3) (%) T = 1.0 C = 1.1 Moderate wasting (WHZ ≥ –3 and < –2) (%) T = 13.4 C = 14.3				

(continued on next page)

APPENDIX D *(continued)*

Study	Country—Program[a]	Intervention[b]	Evaluation Method[c]	Baseline	Findings[d]	Heterogeneity of impacts: Income/poverty level	Heterogeneity of impacts: Maternal education	Heterogeneity of impacts: Other	Remarks
Kazianga, de Walque, and Alderman (2009)	Burkina Faso	F, THR	R, DID	<u>WHZ all children</u> School meals = –0.786 Take-home rations = –1.125 Control = –0.903 <u>Wasting (% of children)</u> All children School meals = 29.5 Take-home rations = 32.3 Control = 42.7	*Children's age: 6–60 months* <u>WHZ</u> *School meals* 6–60 months = 0.005 (n.s.) 12–60 months = 0.062 (n.s.) *Take-home rations* 6–60 months = 0.291 (n.s.) 12–60 months = 0.333*				
Leroy and others (2008)	Mexico—*Oportunidades*	CT, F, G, M, NE, P, T	PSM, DID	<u>WHZ</u> T = 0.30 C = 0.33	*Children's age: 0–24 months* <u>WHZ gain for all age groups</u> 0.085 (ns) <u>WHZ gain by age</u> 0–6 months = 0.465** 6–12 months = –0.172 (n.s.) 12–24 months = 0.103 (n.s.)	<u>WHZ gain (in terciles [T])</u> T1 = 0.118 (ns) T2 = 0.074 (ns) T3 = –0.164 (ns)			Income/poverty levels given in tercile (T1–T3); T1 is the poorest.
Maluccio and Flores (2005)	Nicaragua—RPS	CT, F, G, M, NE, P, T	R, DID	<u>Wasting (%)</u> T = 0.8 C = 0.4	*Children's age: 0–60 months*. <u>Wasting (%)</u> T = 0.2 C = 0.2 Wasting declined in both T and C groups. Project impact is n.s.			No differential effects by gender.	As shown by baseline and follow-up values, wasting is not a problem in the intervention area.
Masanja and others (2005)	Tanzania—IMCI	n.a.	Compared trends and inequalities of health for children 24–59 months of age using two surveys	<u>Wasting (%) in 1999</u> T = 13 C = 11	*Children's age: 12–23 months* <u>Wasting (%) in 2002</u> T = 7 C = 5				The authors compared trends and inequalities of health using surveys from 1999 (n = 2006) and 2002 (n = 1,924) for children under 5 years of age.

Penny and others (2005)	Peru	NE	R		*Children's age: 0–18 months* WHZ at 18 months T = 0.15 C = 0.05 Adjusted difference = 0.048 (*p* = 0.609)				Birth cohort followed from birth to age 18 months.
Quisumbing (2003)	Ethiopia—Food Aid	FD, FFW	Heckman maximum likelihood estimate	n.a.	*Children's age: under 9 years* WHZ Children <5 years in low-asset households FFW has impact; FD and all (FD + FFW) do not. Children 5–8 years in low-asset households Separately, FFW and FD have no impact on WHZ. However, the total value (FFW + FD) has lagged impact. Children <5 years in high-asset households FFW has no impact. FD as well as total value (FD + FFW) have lagged impact. Children 5–8 years in high-asset households FFW has no impact. FD as well as total value (FD + FFW) have lagged impact.	Impact reported by low- and high-asset households (see column to the left).		1. Gender of child. Children <5 years in low-asset households. FFW has positive impact on boys but negative impact on girls; FD has no differential impact; the total value (FD + FFW) has no differential impact either. In high-asset households, FFW has lagged negative impact on girls. FD has lagged impact. No differential impact is observed when the total value of food aid (FFW + FD) is considered. Children 5–8 years: No differential impact by gender on any specifications in low-asset households. Lagged impact in FD and total value (FD + FFW) on girls. 2. Gender of recipient Children <5 years: gender of the child interacted with the gender of the recipient (Girls interacted with female recipient) not significant in low- and high-asset households. Children 5–8 years: interaction term is insignificant and inconsistent (positive and negative impact) in both low- and high-asset households.	

(continued on next page)

Study	Country—Program[a]	Intervention[b]	Evaluation Method[c]	Baseline	Findings[d]	Heterogeneity of impacts			Remarks
						Income/ poverty level	Maternal education	Other	
Ruel and others (2008)	Haiti	F, G, M, NE, P	R	WHZ Preventive = –0.18 Recuperative = –0.18	*Children's age: 12–41 months* WHZ Preventive = –0.22 Recuperative = –0.46 At the end of the three-year intervention, children from preventive communities had significantly higher mean z scores WHZ (0.24) than the recuperative group By age group, significant (higher for preventive group) for age group 12–23 months and 24–35 months, but not for 36–41 months. Wasting Significantly lower (by 4 percentage points) in the preventive model.				Coefficients by age are not reported.
Santos and others (2001)	Brazil	NE	R	WHZ All: T = 0.37 C = 0.55	*Children's age: 0–18 months* WHZ All: 0.09 (n.s.) By age group 0–6 months = –0.02 (n.s.) 6–12 months = 0.24 (n.s.) 12–18 months = 0.28**				
Schipani and others (2002)	Thailand	Mixed gardening	Matching program and nonprogram areas	n.a.	*Children's age: 12–84 months* No significant impact on wasting (WHZ ≤ –2): Rainy season T = 10.0 C = 6.6 Cool season T = 0.0 C = 0.0 Hot season T = 3.7 C = 0.0				No baseline information is used in the analysis. The study is based on single difference from a cross-sectional data collected in three different seasons.

Schroeder and others (2002)	Vietnam—CENP	D, F, G, NE	R, multivariate	WHZ T = –0.66 C = –0.90 Wasting (%) (WHZ <–2 SD) T = 1.7 C = 10.1	*Children's age: 5–30 months* WHZ T = 1.25 C = –1.35 The difference is not significant ($p = 0.12$). Wasting (%) T = 14.9 C = 12.7 The difference is not significant ($p = 0.63$).				Unbalanced baseline. The difference is significant. T better off. Small number of observations in the T at baseline for wasting.
White and Masset (2007)/IEG (2005)	Bangladesh—BINP	F, G, M, NE, P	PSM		*Children's age: 6–23 months* WHZ. Midterm: 0.10*** Endline: 0.03 (n.s.)				Average treatment effect, on the treated, one-to-one matching (IEG 2005, p. 170)

Source: IEG analysis.

Notes: C = control group; T = treatment group; WHZ = weight-for-height z-score; n.a. = not available or not mentioned in the reviewed material; n.s. = not statistically significant. Statistical significance: * = p < =.10, ** = p <=.05, *** = p <=.01.

a. BINP = Bangladesh Integrated Nutrition Project; CENP = Community Empowerment and Nutrition Project; ECD = early childhood development program; IMCI = Integrated Management of Childhood Illness; RPS = *Red de Protección Social*.

b. CT = cash transfer; DC= day care; D = de-worming; F = feeding; FD = free food distribution; FFW = food for work; G = growth monitoring; M = micronutrient supplement , NE = nutrition education; P = prenatal services; T = treatment of illness; THR = take-home rations.

c. DID = difference in difference/double difference; PSM = propensity score matching; R = randomized.

d. kg = kilogram; IMCI = Integrated Management of Childhood Illness

APPENDIX E

Impact Evaluations of Birthweight and Low Birthweight

Study	Country—Program	Intervention[a]	Evaluation Method[b]	Impacts[c]	Heterogeneity of impacts	Remarks
Attanasio and others (2005)	Colombia—*Familias en Acción*	CT, NE	PSM, DID	Significant impact on weight gain in urban areas.	Weight gain by newborns in urban areas is 0.578 kg (significant) and in rural areas 0.176 kg (n.s.).	The program does not require pregnant women to visit clinics to obtain prenatal care. Attendance at meetings at which health, hygiene, and nutrition issues are discussed are not compulsory, although the program encourages women to attend such meetings.
Barber and Gertler (2008)	Mexico—*Oportunidades*	CT, F, M, NE, P	IV	Beneficiaries had 127.3 g (95% CI: 21.3, 233.1) higher birthweight among participating women and a 4.6 percentage point reduction in LBW. Program impact using the average beneficiary time on program amounts to 68.3 g ($p = 0.05$) and program impact from cash received amounts to 78.2 g ($p = .07$).		
Christian and others (2003)	Nepal	M	R	Folic acid-iron increased mean birthweight by 37 g (95% CI: 16, 90) and reduced the percentage of LBW babies (< 2,500 g) from 43% to 34% (16%; relative risk = 0.84, 0.72–0.99). Multiple micronutrient supplementation increased birthweight by 64 g (12–115 g) and reduced the percentage of LBW babies by 14% (0.86, 0.74 to 0.99).		Folic acid–iron and multiple micronutrients increased head and chest circumference of babies, but not length.
Friis and others (2004)	Zimbabwe	M	R	Micronutrient supplementation was associated with increased birthweight, 49 g (95% CI: –6 to 104 g; $p < 0.08$), but was not associated with LBW.	No difference by human immunodeficiency virus status of women.	
Gupta and others (2007)	India	M	R	Incidence of LBW reduced in the treatment group. No difference in birthweight.		Facility-based randomized trial on malnourished pregnant women. The T (micronutrient) received 29 vitamins and minerals once a day for 37–77 days; the C received placebo for 15–66 days. All participants received diet advice; usual prenatal services, including immunization; and iron and folic acid supplements.

Iannotti and others (2008)	Peru	M	R	LBW = 2.2% in both the T and C groups. No significant difference on weight at birth.		The T received zinc, iron and folic acid; the C received iron and folic acid.
Macours, Schady, and Vakis (2008)	Nicaragua—*Atención a Crisis*	CT, F, M, NE, P	R	No significant program effect on LBW. Treatment effect (weight gain) for 0–5-month-old infants was 0.161 kg (insignificant). Mean birthweight in the control group was 2.987 kg.		
Menéndez and others (2008)	Mozambique	Sulphadoxine-pyrimethamine with insecticide-treated nets.	R	No significant impact on LBW for the whole sample, but significant in subsamples (see column at right).	Impacts on LBW by HIV status and gravidity. No change by HIV status, but by gravidity, incidence of LBW was lower among women with four or more prior pregnancies.	All participants in the T and C groups received long lasting insecticide-treated nets. Those in the T received two doses sulphadoxine-pyrimethamine. Pregnant women with HIV infection were not given sulphadoxine-pyrimethamine.
Osrin and others (2005)	Nepal	M	R	Significant impact on birthweight. The mean (±SD) of T and the C groups are 2.810 (± 0.453) and 2.733 (± 0.422), respectively. Mean difference = 77g (95% CI 24–130; p = 0.004). Fall (25%) in the incidence of LBW.		The T received multiple micronutrient supplementations (recommended daily allowance of 15 vitamin and minerals) with iron and folic acid; the C received iron and folic acid only.
Ramakrishnan and others (2003)	Mexico	M	R	No significant difference between the T and C groups in birthweight or incidence of LBW. The mean (±SD) of the T and C groups are 2.981 (± 0.391) and 2.977 (± 0.393), respectively.		The T group received iron and multiple micronutrient supplementation; the C received iron-only supplementation
Zeng and others (2008)	China	M	R	Micronutrients had significant but modest impact on birthweight. Birthweight was 42 g (95%CI 7–78 g) higher in the multiple micronutrient group than in the folic acid group. No significant difference between the T and C groups for LBW.		Participants were randomized into three groups. The C received folic acid alone. In the two T groups, one received daily folic acid + iron, and the other received multiple micronutrient with recommended daily allowance of 15 vitamin and minerals.

Source: IEG analysis.
Note: C = control group; g =gram; HIV = human immunodeficiency virus; kg = kilogram; LBW = low birthweight; T = treatment group; n.s = not significant.
a. CT = cash transfer; F = feeding; M = micronutrient supplement; NE = nutrition education; P = prenatal services.
b. DID = difference-in-difference/double difference; IV = instrumental variable/2SLS; PSM = propensity score matching; R = randomized.
c. CI = confidence interval; SD = standard deviation

APPENDIX F

Impact Evaluation Basics

Program *impact* in this review is defined as the difference in child anthropometric outcomes of two statistically comparable groups—one with the program (the treatment group) and the other without it (the control group). The magnitude of impact can be either an *intent-to-treat* or a *treatment-on-the treated* estimate. The average intent-to-treat effect is an estimate of the average impact of the availability of the program on eligible beneficiaries in treatment areas, whether or not they were actually treated. Including the untreated in the treatment group may bias the results downward.

In contrast, the average treatment-on-the treated parameter is the effect of the program on those who actually received the treatment. The intent-to-treat estimate can be a parameter of interest in nutrition impact evaluations. For example, a cost-effectiveness analysis of a school-based deworming or supplementation program needs to consider the fact that all children may not be at school on the day of the treatment and that tracking children at home may not be practical. Therefore, in this case, the parameter of interest is intent-to-treat (Duflo and others 2007). There are many cases where other data-related and methodological concerns (mainly self-selection into the program) make using intent-to-treat estimations better than the treatment-on-the treated effect.

Experimental or *randomized design* is regarded as the most robust of impact evaluation methodologies. Because the beneficiaries of a program cannot be both receiving and not receiving it, the control group must be constructed from a group that is very similar. One critical difference between a reliable and an unreliable impact evaluation, therefore, is how well this counterfactual approximates the treatment group in the absence of the intervention. Random assignment to the program ensures initial equivalence of the beneficiary (treatment) and nonbeneficiary (control or comparison) groups. It implies that both observable and unobservable characteristics in the two groups are statistically identical. In that case, the impact of the program is measured by the difference in mean outcomes between the treatment and the control groups. In addition to this simplicity in interpreting and conveying the results, a randomized evaluation design eliminates the possibility that specification error is influencing the results (Duflo and Kremer 2003; Duflo and others 2007). In this review, the primary identification strategy of 21 evaluations (46 percent of those reviewed) is based on randomization.

It is important to note that in practice, particularly in development applications, randomization can be difficult to implement (Baker 2000; Ravallion 2009a). First, it may not be ethical to deny treatment to otherwise eligible individuals or to provide treatment to those who do not need it. Second, it is not always politically possible to provide treatment to one group and to deny or delay treatment to another. Third, not all interventions are amenable to randomized evaluation. For example, some interventions are conducted at the national level, and the scope may mean that there is no possibility for randomization. Fourth, results could be invalidated or contaminated as a result of spillovers and changes in the behavior of individuals in the treatment group or the control group. Fifth, the generalizability (external validity) of the results may be a source of concern. Sixth, randomized designs can be expensive and time consuming.

Proponents of randomization challenge some of these limitations (Duflo and others 2007). For example, on ethics, it is argued that it would be wrong "to assume that one would be denying the poor a beneficial intervention until an idea has been properly evaluated" (World Bank 2007c). Moreover, other ethical and political issues can be addressed by extending the program in the control areas at a later stage and by selecting the treatment and control groups in a politically transparent manner (Baker 2000).

Concerning contamination, Duflo and others (2007) argue that spillover effects can be captured if randomization occurs at a higher level. For example, Miguel and Kremer (2004) randomized at the school level and found larger effects of de-worming drugs than other evaluations did based on individual-level randomization. Regarding costs, Duflo and Kremer (2003) argue that evaluation costs can be reduced by conducting a series of evaluations in the same area. Finally, problems of external validity also apply to nonexperimental methods.

Quasi-experimental designs comprise a **class of causal evaluation designs that define a control group through some nonrandom process. The identification strategy in 25 of the 46 reviewed evaluations (54 percent) is based on these nonrandom processes.** Econometric techniques are used to generate comparison groups that resemble the treatment group, at least in observed characteristics. Among the advantages of these approaches are that they can use existing data and are cheaper and quicker to implement. However, one critical problem with quasi-experimental approaches is selection bias. Randomization balances the selection bias between the treated and the untreated samples (Heckman and Smith 1995), but nonrandomized approaches use complex methods to correct it. Quasi-methods include matching techniques, difference-in-difference (DID) or double-difference methods, instrumental variables methods, regression discontinuity, and reflexive comparisons.

The following methods were used by one or more of the reviewed studies.

- *Matching methods or constructed controls*—The main task is to pick an ideal comparison group that matches the treatment group. The most widely used type of matching is propensity score matching (PSM), in which the comparison group is matched to the treatment group on the basis of a set of observed characteristics. In this method, treated and untreated cases are matched on the basis of propensity scores (the predicted probability of participating in the intervention, given observed characteristics). The closer the score, the better the match.

  However, PSM can introduce error if the treated and the untreated groups do not have substantial overlap in observed characteristics. For example, PSM would lead to regression toward the mean if the worst cases of the untreated were compared with the best cases of the treated group. Other drawbacks of PSM and other matching methods include the need for large samples, the strong assumption that individuals in the matched control group did not choose to be untreated, and hidden bias that might remain because of differences between the treated and the untreated groups in unobservable characteristics. For example, in the *Hogares Comunitarios* program, Attanasio and Vera-Hernandez (2004) show that PSM would show counterintuitive results on the impact of the program. They argue that a comparison of attending and nonattending children based on observables alone would be misleading "as it ignores the endogeneity of the participation decisions." In this review, 12 of the 46 evaluations (26 percent) used PSM.

- *Double difference* or *DID*—This method compares the treatment and control groups (first difference) before and after the intervention (second difference). The validity of this analysis depends on the assumption on the parallel evolution of the outcome in the absence of the treatment. Eleven evaluations (24 percent) reviewed for this study used DID in combination with other methods.

- *Instrumental variables*—The instrumental variables method recognizes that program placement is not random, but purposive. Therefore, this method identifies the exogenous component of the variance in program placement by using instrumental variables that matter to participation to the program but not to outcomes, given participation. The validity of this method depends on the quality of the instrument. The instrumental variables method was used in six evaluations (13 percent) reviewed for this study.

Endnotes

Chapter 1

1. Estimates are for 2005.
2. De Onis and Blössner (2000), based on an analysis of 160 national surveys from 94 countries. *Overweight* is defined as a weight that is more than two standard deviations above that of the reference population for a given height. Among the regions with the highest rates of overweight are Northern Africa (8.1 percent), Southern Africa (6.5 percent), and Latin America and the Caribbean (4.4 percent).
3. This is the share of the lending portfolio managed by the Health, Nutrition, and Population Sector with nutrition objectives; the share of projects managed by other sectors that have nutrition objectives or components was not quantified.
4. The renewed commitment is evidenced in part by the recent recruitment of six nutrition specialists to address malnutrition, particularly in Africa and South Asia.
5. More recently, the Poverty Reduction and Economic Management Network issued a handbook entitled *Methodologies to Evaluate the Impact of Large-Scale Nutrition Projects*.
6. Because of this complexity, Bhutta and others (2008) note that "the choice [of intervention] will depend on the actual nature and distribution of the malnutrition problem, its causes, and the type of resources that are available" (p. ix).
7. As an exception, in China and Madagascar, where the edible-salt industry is concentrated in a few producers, salt iodization can be nearly universalized and little choice is exercised by households (Goh 2001).
8. The conclusions on breastfeeding promotion, complementary feeding, and food supplementation in populations with and without sufficient food, for example, were based on 10 studies—3 in food-secure populations (defined as having average income of more than $1/day) and 7 in non-food-secure populations (Bhutta and others 2008).
9. Despite the lack of data on the effectiveness of large-scale interventions, the authors nevertheless classify a relatively long list of specific nutrition interventions into four categories as the basis for their recommendations on scaling up: (a) interventions for which "evidence was sufficiently robust to recommend their use in most countries with high burdens of undernutrition"; (b) those that might be recommended for countries in specific situational contexts; (c) those with insufficient or variable evidence; and (d) those for which the evidence showed little or no impact. These recommendations are summarized in appendix A.
10. It is important to note that a primary objective of CCTs is to affect poverty, as well as human development outcomes such as nutrition.
11. This may be due in part or mostly to a failure by the studies themselves to examine the heterogeneity of impacts (Heckman and Smith 1995; Ravallion 2009). However, nutrition impact evaluations often do present results across different age groups—the main exception.
12. Bhutta and others (2008) highlight this evidence on "effectiveness and cost-effectiveness of nutritional interventions in national health systems, single and packaged, for impact on stunting and weight gain."
13. In this regard, it is important to note that child anthropometric outcomes were often not the only outcomes anticipated from these interventions. A comparative assessment of interventions across their other major objectives (both in terms of other nutritional outcomes, as well as cognitive and poverty reduction outcomes) is beyond the scope of this paper.

Chapter 2

1. Studies of the impact of interventions on other anthropometric outcomes, such as upper-arm circumference and skinfold thickness, were excluded.
2. Most of these evaluations measured program effects of the interventions on several other schooling and health outcomes. Further, for some of the interventions (such as CCTs and micronutrient interventions), improving anthropometric outcomes was not the primary objective. Interventions with little impact on anthropometric outcomes might have significant impacts on these other primary outcomes; however, these are not reviewed here.
3. CCTs and UCTs, for example, are generally offered to low-income households.
4. These evaluations nevertheless often control for demographic and socioeconomic characteristics to reduce idiosyncratic variation and to improve the power of the estimates (for example, Bobonis, Miguel, and Sharma 2006; Gertler 2004; Morris and others 2004; Paxson and Schady, forthcoming).
5. Quasi-experimental methods may be adopted when randomization fails to equate the treatment and control or when no baseline information is available.

6. Macours and others (2008) studied impacts on children 0–23, 24–47, and 48–71 months old; Maluccio and Flores (2005) studied children age 0–60 months.

7. Morris and others (2004) found no impacts on children 0–23, 24–47, and 48–84 months old.

8. Agüero and others (2007) consider the first three years of life as a "nutritional window" vital for larger program impact. They argue that a treatment that covers much of the child's early age boosts the HAZ, and there are no gains for treatments covering less than 20 percent of the child's nutritional window.

9. The program included a behavior change and communication component. The preventive model targeted all children age 6–23 months, and the recuperative model targeted underweight children age 6–60 months

10. In fact, FFW had a negative impact on HAZ for children under five and for girls five to nine years of age in low-asset households ($p < 0.10$ and $p < 0.05$, respectively). Lagged food distribution had a negative impact on HAZ for children five to nine years of age in high-asset households ($p < .001$), although the magnitude of the impact is quite small.

11. Participation in the program, captured by the current attendance measure, was associated with an increase in the HAZ by 0.486, which is equivalent to 2.36 centimeters in height for a boy or 2.39 centimeters for a girl at age 72 months. The exposure model suggests that impact increases when participation is adjusted by age. The age-adjusted increase is 0.78 in HAZ, which is equivalent to a 3.78-centimeter increase in height for a boy or 3.83-centimeter increase for a girl 72 months old.

12. The finding was statistically significant at the $p = 0.10$ level.

13. Das Gupta and others (2005) (ICDS, for children age 0–3 or 0–4 years), Schipani and others (2002) (gardening, for children age 1–7 years).

14. Of 15 coefficients representing children of different ages and exposures to the program, 3 indicated a significant reduction in stunting and 5 indicated an increase. The remaining 7 coefficients were insignificant.

15. The authors speculate that this counterintuitive result might be caused by a perception by beneficiaries that "benefits would be discontinued if the child started to grow well."

16. The evaluations in Kenya and India had similar designs and found impacts on other educational and health outcomes. However, in India the program raised WAZ but not HAZ for children age 2–6, while the opposite was the case in Kenya for children aged 6–18.

17. However, the comparability of the program and nonprogram areas was not well established. The subsequent evaluation by White and Masset (2007) with a more rigorous methodology that used propensity score matching did not report results on underweight.

18. At baseline in 2000, 13.7 percent and 14.3 percent of the children in the program and nonprogram areas, respectively, were underweight, respectively, with just a –0.6 insignificant difference between them. The net underweight averted by the program was 5.5 percentage points.

19. For example, the average regional prevalence of stunting, underweight, and wasting for 2000–07 based on the National Center for Health Statistics reference population is as follows: Sub-Saharan Africa (38 percent, 28 percent, and 9 percent); Latin America and the Caribbean (16 percent, 6 percent, and 2 percent); and South Asia (46 percent, 45 percent, and 18 percent). http://www.childinfo.org/index.html.

20. Of the 15 results, 7 had significant and positive impacts.

21. The prevalence of wasting in Haiti is for children younger than 0–59 months in 2000. http://www.childinfo.org/undernutrition_wasting.php.

22. However, their control areas were less than ideal. Unfortunately, White and Masset (2007) did not report findings on wasting using more robust PSM techniques.

23. All in all, they report 15 results, with 9 showing impact. Of the 9, 6 were with their expected negative signs.

24. Only the evaluation of Colombia's CCT, *Familias en Acción,* by Attanasio and others 2005 used a quasi-experimental design (PSM and difference-in-difference techniques).

25. The average beneficiary time in the CCT program contributes 68 grams, and the amount of cash received is associated with a 78.2-gram weight gain. Program time measures the number of months between the date of receipt of the first cash transfer and the date of birth.

26. The sample size (including treatment and control) for this part of the analysis is 174. The authors suggest that lack of significant impact might be due to the small size of the sample.

27. These impacts were not found for all women (just for this subgroup), although the evaluation did find impacts on malaria and anemia.

28. The "better-off" communities were the third of communities with the lowest incidence of poverty.

29. In fact, table 6 of Quisumbing (2003) shows that FFW improves the WHZ of boys under five in low-asset households and worsens the WHZ of girls.

30. The other intermediate outcomes measured were pregnancy knowledge and practice (three evaluations) and hygiene behavior (one evaluation).

31. The cost of a de-worming program per pupil per year is \$0.49, and the authors show that 99 percent of the reduction in DALYs was attributable to the averted schistosomiasis.

32. A scenario that is taken into consideration is a preschool program that results in a 2 percent increase in height at childhood, a 5 percent increase in cognitive skills and a one-year increase in grades completed, and a corresponding

one-year increase in the age of school completion. However, the program did not improve child nutritional status.

33. They estimate program impact on WAZ of children and show that gains are larger for more educated mothers for villages with better infrastructure.

34. Hossain and others (2005) also found an increase in knowledge in project areas, compared with nonproject areas, but concluded that there was no impact on child nutrition outcomes. However, the project and nonproject areas may not have been comparable.

35. See in particular the DHS evidence presented in appendix E. However, the surveys used for the impact evaluation did not include these measures, so it was not possible to examine BINP impacts for women who did and did not face these constraints.

Chapter 3

1. The Development Impact Evaluation Initiative (DIME) is a Bank-wide collaboration involving thematic networks, Regional units, and the research group under the guidance of the World Bank's Chief Economist. There are 27 completed or ongoing evaluations reported on the DIME Web site that measure impacts on anthropometric outcomes, 6 of which are reviewed in this study. Of the 21 remaining, two-thirds measure the impact of health or nutrition interventions, and a third measure the impact of social protection interventions (CCTs, social funds). About half involve a randomized design, a quarter used a quasi-experimental design, three use both methods, and for three the methodology was not reported. A third are in Sub-Saharan Africa, and a quarter each are in Latin America and the Caribbean and South Asia; none measures nutrition outcomes in Eastern Europe and Central Asia or the Middle East and North Africa. More than half of these nutrition impact evaluations are linked to World Bank projects. Six have been completed. The Spanish Impact Evaluation Fund directly funds impact evaluations, preferring experimental designs, but to date it has funded none of the proposals on nutrition.

2. IEG was able to interview project team leaders and evaluators for all eight programs; policy makers were interviewed for six of the eight countries (Bolivia and the Philippines were not reached).

3. The conditionality was announced but never enforced. So for all intents and purposes, the program was an unconditional transfer.

4. *Hogares de Bienestar Infantil,* in Colombia, had been evaluated in 1992 and was found to be successful (World Bank 1993, p. 14).

5. Projects in Colombia, Ecuador, Bangladesh, and the Philippines incorporated impact evaluations explicitly in the PAD; the other projects all called for baseline, midterm, and endline surveys or evaluations.

6. In fact, many different evaluative activities were programmed into the Bolivia project.

7. The triggers included selection of a firm for the baseline survey of BDH; a methodology and implementation schedule (first loan); adequate progress in implementation of the evaluation, according to the plan (second loan); and changes in the design, budget, and implementation of BDH based on the results of the impact evaluation (third loan) (World Bank 2003a).

8. Karim and others (2003) measured the impact of the project as the difference in outcomes between the baseline and endline surveys in project areas; there was no attempt to compare results with nonproject areas.

9. The consortium included *Econometria Consultores;* the Institute for Fiscal Studies at University College, London; and *Sistemas Especializados de Informacion.*

10. The non-Bank researchers involved in the evaluation of *Familias en Acción* and BINP were not involved in the design of the projects they evaluated.

11. The Office of Population Studies, San Carlos University, Cebu, Philippines, and the Institute of Public Health at Makerere University, Kampala, Uganda. The White and Masset (2007) evaluation of BINP used existing data sets; the evaluation by Hossain and others (2005) financed their own data collection, but it is unclear which organization collected the data.

12. The evaluation was nevertheless part of the policy matrix for the First Programmatic Human Development Reform Project.

13. The two research proposals and funding were for community nutrition program impact evaluations in Madagascar and Senegal (Alderman and Rokx 2003, request for $207,200) and for evaluation of the three early child development programs in Bolivia, the Philippines, and Uganda (Alderman and van der Gaag circa 1997, request for $395,500).

14. The impact evaluation of BDH in Ecuador received $400,000 from the Spanish Impact Evaluation Fund and a $1 million grant from the Japanese Trust Fund; additional data collection by Galasso and Umapathi (2009) of community nutrition in Madagascar was funded with grants from the Bank–Netherlands Partnership Program and UNICEF; the evaluation of BINP and other maternal and child health programs by IEG was supported by $230,000 from a Department for International Development partnership and $23,400 from a Danish trust fund.

15. The IEG budget supported the BINP evaluation (which was combined with the evaluation of several other maternal and child health programs) to the sum of $165,625.

16. The NGOs had launched sensitization and mobilization activities in the communities before the impact evaluation

design was finalized, putting them in an awkward position vis-à-vis communities previously mobilized for which implementation would have to be deferred.

17. The main difference between the treatment and control municipalities was that the controls lacked a bank, which was essential for processing the transfer.

18. Orazio Attanasio, personal communication.

19. Alderman (2007), Armecin and others (2006), Linnemayr and Alderman (2008), Behrman, Cheng, and Todd (2004), Galasso and Umapathi (2009), Galasso and others (2009), White and Masset (2007). Matching methods also have limitations, however. It is possible to match only on the basis of characteristics that are observed in both the treatment and control populations.

20. However, it is important to control for the characteristics of the communities or individuals enlisted at different times. For example, the program may have initially targeted the neediest individuals or communities.

21. The cash transfer and early child development interventions often aimed to affect other outcomes, including educational attainment and cognitive outcomes, and, in some cases, other health outcomes. However, this section focuses narrowly on the findings on child anthropometric status.

22. The authors point to cultural factors—the lack of control of women in decisions regarding food purchase and preparation—as possibly explaining the fact that better knowledge does not seem to have led to much better outcomes.

23. In the Bolivia early child development project, food accounted for about half of the total project cost of $36/child/month.

24. This perhaps is not surprising, given the short implementation period (18 months) and the well-documented finding in the literature that the weight and height of children under two are particularly sensitive to nutritional inputs.

25. Almost all of the studies examined the impacts across different age groups of children (the exception being the de-worming evaluation in Uganda). Here we review heterogeneity in impacts across socioeconomic characteristics and access to services. The evaluation of BINP by White and Masset presented results on the heterogeneity of intermediate behavioral outcomes but not nutrition impacts.

26. Note, however, that this result does not apply to HAZ individually but rather to a synthetic index of three "physical" outcome measures that included HAZ.

27. The evaluations of the cash transfer programs in Colombia and Ecuador are among those that did not examine impacts as a function of the availability of public services. Yet low access to health care conceivably could be a reason for nonparticipation or nonadherence in the Colombia CCT program, and, in the case of Ecuador, the availability and quality of health services is likely to affect the extent to which additional cash income is translated into health outcomes.

28. The estimate of $43 is attributed by Behrman, Cheng, and Todd (2004) to Ruiz (1996). The Implementation Completion and Results Report for the project put the cost at $30/month/child initially, which was brought down to $22/month/child. Subsequent changes to the program (after the impact evaluation) brought the cost down to $2/month/child, based on eight months of implementation.

29. One of the difficulties in conducting cost-benefit analysis is that there is often no country-specific data on how nutritional and other impacts from the program affect long-run earnings, on the basis of which to calculate the benefits. Thus, they are often extrapolated from studies in other settings.

30. The authors calculate, according to simulations (not based on the impact evaluation parameters), that the cost of preventing one case of underweight by simply financing a rice ration would be on the order of $110 per year and the cost per life saved $2,223.

31. IEG was unable to interview policy makers from Bolivia and the Philippines.

32. Cited in the Implementation Completion and Results Report. In retrospect, it is fortunate that some of the municipalities in the impact evaluation baseline survey had already been enlisted into the program. Had that not been the case, there would have been no quick evidence that the program was effective to provide to the new government. It was also reported by informants that evidence from the Progresa evaluation in Mexico was influential in the decision to continue the program.

33. According to informants, the cash transfers for rural families in the most vulnerable municipalities are conditioned on the number of annual visits for children under two on their "healthy child card" and on the weight register at the health facility. Children under one year of age must show at least six visits, and children between one and two years must show at least three visits.

Chapter 4

1. This point is also made in a 2008 letter to the editor of *The Lancet,* in which Shekar and 17 signatories highlight the need to expand the research agenda to include the "delivery science" to "understand implementation and cost effectiveness at scale" of nutrition interventions.

Bibliography

*Agüero, Jorge M., Michael Carter, and Ingrid Woolard. 2007. "The Impact of Unconditional Cash Transfers on Nutrition: The South African Child Support Grant." International Poverty Centre Working Paper No. 39. International Poverty Centre, United Nations Development Programme, Brasilia, Brazil.

*Alderman, Harold. 2007. "Improving Nutrition through Community Growth Promotion: Longitudinal Study of the Nutrition and Early Child Development Program in Uganda." *World Development* 35 (8): 1376–89.

Alderman, Harold, and Jere R. Behrman. 2006. "Reducing the Incidence of Low Birthweight in Low-Income Countries Has Substantial Economic Benefits." *World Bank Research Observer* 21 (1): 25–48.

*Alderman, Harold, Hans Hoogeveen, and Mariacristina Rossi. 2006. "Reducing Child Malnutrition in Tanzania: Combined Effects of Income Growth and Program Interventions." *Economics and Human Biology* 4: 1–23.

*Alderman, Harold, Joseph Konde-Lule, Isaac Sebuliba, Donald Bundy, and Andrew Hall. 2006. "Effect on Weight Gain of Routinely Giving Albendazole to Preschool Children during Child Health Days in Uganda: Cluster Randomized Controlled Trial." *British Medical Journal* 333: 122–26.

*Alderman, Harold, Brian Ndiaye, Sebastian Linnemayr, Abdoulaye Ka, Claudia Rokx, Khadidiatou Dieng, and Menno Mulder-Sibanda. 2009. "Effectiveness of a Community-Based Intervention to Improve Nutrition in Young Children in Senegal: A Difference-in-Difference Analysis." *Public Health Nutrition* 12 (5): 667–73.

Alderman, Harold, and Claudia Rokx. 2003. "Community Nutrition: Evaluation of Impacts in Africa." Request for Research Support Budget Grant: Administrative Summary. June 30.

Alderman, Harold, and Elizabeth King. Undated. "Evaluation of the Impact of Investments in Early Child Development on Nutrition and Cognitive Development." World Bank Research Committee Project Completion Report No. P052067, Washington, DC.

Alderman, Harold, and Jacques van der Gaag. Undated, circa 1997. "Evaluation of the Impact of Investments in Early Child Development on Nutrition and Cognitive Development: A Research Proposal." Research Proposal Submitted to the World Bank Research Committee.

Allen, Lindsay H., and Stuart R. Gillespie. 2001. "What Works? A Review of the Efficacy and Effectiveness of Nutrition Interventions." ACC/SCN Nutrition Policy Paper 19; Asian Development Bank (ADB) Nutrition and Development Series 5. United Nations Administrative Committee on Coordination/Sub-Committee on Nutrition and the Asian Development Bank, Geneva and Manila.

*Armecin, Graeme, Jere R. Behrman, Paulita Duazo, Sharon Ghuman, Socorro Gultiano, Elizabeth M. King, and Nanette Lee. 2006. "Early Childhood Development Programs through an Integrated Program: Evidence from the Philippines." Policy Research Working Paper 3922, World Bank, Washington DC.

*Attanasio, Orazio, Luis Carlos Gómez, Patricia Heredi, and Marcos Vera-Hernández. 2005. "The Short-term Impact of a Conditional Cash Subsidy on Child Health and Nutrition in Colombia." Institute of Fiscal Studies, London.

*Attanasio, Orazio, and M. Vera-Hernández. 2004. "Medium and Long-Run Effects of Nutrition and Child Care: Evaluation of a Community Nursery Program in Rural Colombia." IFS Working Paper EWP04/06, Institute for Fiscal Studies, London.

Baker, J. 2000. *Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners*. Washington, DC: World Bank.

*Barber, Sarah L., and Paul J. Gertler. 2008. "The Impact of Mexico's Conditional Cash Transfer Programme, Oportunidades, on Birthweight." *Tropical Medicine & International Health* 13 (11): 1405–14.

*Behrman, Jere R., Yingemei Cheng, and Petra E. Todd. 2004. "Evaluating Preschool Programs When Length of Exposure to the Program Varies: A Nonparametric Approach." *Review of Economics and Statistics* 86 (1): 108–32.

*Behrman, Jere R., and John Hoddinott. 2005 "Programme Evaluation with Unobserved Heterogeneity and Selective Implementation: The Mexican PROGRESA Impact on Child Nutrition." *Oxford Bulletin of Economics and Statistics* 67 (4): 547–69.

Bhutta, Zulfiqar A., Tahmeed Ahmed, Robert E. Black, Simon Cousens, Kathryn Dewey, Elsa Giugliani, Batool A. Haider, Betty Kirkwood, Saul S. Morris, H. P. S. Sachdev,

*These references are the 46 impact evaluations that the Independent Evaluation Group reviewed for this study.

and Meera Shekar, for the Maternal and Child Undernutrition Study Group. 2008. "What Works? Interventions for Maternal and Child Undernutrition and Survival." *The Lancet* 371 (9610): 417–40.

Black, Robert E., Lindsay H. Allen, Zulfiqar A. Bhutta, Laura E. Caulfield, Mercedes de Onis, Majid Ezzati , Colin Mathers, and Juan Rivera. 2008. "Maternal and Child Undernutrition: Global and Regional Exposures and Health Consequences." *The Lancet* 371: 243–60.

*Bobonis, Gustavo, Edward Miguel, and Charu Puri Sharma. 2006. "Anemia and School Participation." *Journal of Human Resources* 41 (4): 693–721.

Caulfield, L., S. Richard, J. Rivera, P. Musgrove, and R. Black. 2006. "Stunting, Wasting, and Micronutrient Deficiency Disorders." In *Disease Control Priorities in Developing Countries*, ed. Dean T. Jamison, Joel G. Breman, Anthony R. Measham, George Alleyne, Mariam Claeson, David B. Evans, Prabhat Jha, Anne Mills, Philip Musgrove, chapter 28. Washington, DC: World Bank.

*Christian, Parul, Subarna K. Khatry, Joanne Katz, Elizabeth K. Pradhan, Steven C. LeClerq, Sharada Ram Shrestha, Ramesh K. Adhikari, Alfred Sommer, and Keith P. West Jr. 2003. "Effects of Alternative Maternal Micronutrient Supplements on Low Birthweight in Rural Nepal: Double-blind Randomized Community Trial." *British Medical Journal* 326: 571–74.

*Das Gupta, Monica, Michael Lokshin, Michele Gragnolati, and Oleksiy Ivaschenko. 2005. "Improving Child Nutrition Outcomes in India: Can the Integrated Child Development Services Program Be More Effective?" Policy Research Working Paper 3647, World Bank, Washington, DC.

De Onis, Mercedes, and Monika Blossner. 2000. "Prevalence and Trends of Overweight among Preschool Children in Developing Countries." *American Journal of Clinical Nutrition* 72: 1032–39.

Deaton, Angus. 2008. "Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development." The Keynes Lecture in Economics, British Academy, October 9.

Duflo, E., and M. Kremer. 2003. "Use of Randomization in the Evaluation of Development Effectiveness." Paper prepared for the World Bank Independent Evaluation Group, Conference on Evaluation and Development Effectiveness, Washington, DC, July 15–16.

Duflo, E., M. Kremer, and R. Glennerster. 2007. "Using Randomization in Development Economics Research: A Toolkit." Discussion Paper 6059, Center for Economic Policy Research, London.

Fernald, Lia C. H., Paul J. Gertler, and Lynnette M. Neufeld. 2008. "Role of Cash in Conditional Cash Transfer Programmes for Child Health, Growth, and Development: An Analysis of Mexico's Oportunidades." *The Lancet* 371: 828–37.

Fiszbein, Ariel, and Norbert Schady with Francisco H. G. Ferreira, Margaret Grosh, Nial Kelleher, Pedro Olinto, and Emmanuel Skoufias. 2009. *Conditional Cash Transfers: Reducing Present and Future Poverty*. Washington, DC: World Bank.

*Friis, Henrik, Exnevia Gomo, Norman Nyazema, Patricia Ndhlovu, Henrik Krarup, Pernille Kœstel, and Kim F. Michaelsen. 2004. "Effect of Multimicronutrient Supplementation on Gestational Length and Birth Size: A Randomized, Placebo-Controlled, Double-Blind Effectiveness Trial in Zimbabwe." *American Journal of Clinical Nutrition* 80: 178–84.

*Galasso, Emanuela, and Nithin Umapathi. 2009. "Improving Nutritional Status through Behavioral Changes: Lessons from Madagascar." *Journal of Development Effectiveness* 1 (1): 60–85.

*Galasso, Emanuela, and Jeffrey Yau. 2006. "Nutritional Gains from Extended Exposure to a Large-Scale Nutrition Program." Policy Research Working Paper 4508, World Bank, Washington, DC.

*Gertler, Paul. 2004. "Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment." *The American Economic Review* 94 (2): 336–41.

Goh, Chorching. 2001. "An Analysis of Combating Iodine Deficiency: Case Studies of China, Indonesia, and Madagascar." IEG Working Paper Series, No. 18, World Bank Independent Evaluation Group, Washington, DC.

*Gupta, Piyush, Mily Ray, Tarun Dua, Gita Radhakrishnan, Rajeev Kumar and H. P. S. Sachdev. 2007. "Multimicronutrient Supplementation for Undernourished Pregnant Women and Birth Size of Their Offspring." *Archives of Pediatric Adolescent Medicine* 161: 58–64.

Habicht, Jean-Pierre, Reynaldo Martorell, and Juan A. Rivera. 1995. "Nutritional Impact of Supplementation in the INCAP Longitudinal Study: Analytic Strategies and Inferences." *Journal of Nutrition* 125(4): 1042S–50S.

Heckman, James J., and Jeffrey A. Smith. 1995. "Assessing the Case for Social Experiments." *Journal of Economic Perspectives* 9 (2): 85–110.

Hoddinott, John, John A. Maluccio, Jere R. Behrman, Rafael Flores, and D. Reynaldo Martorell. 2008. "Effect of a Nutrition Intervention during Early Childhood on Economic Productivity in Guatemalan Adults." *The Lancet* 371: 411–16.

*Hossain S., M. Moazzem, Arabella Duffield, and Anna Taylor. 2005. "An Evaluation of the Impact of a US$60 Million Nutrition Programme in Bangladesh." *Health Policy and Planning* 20 (1): 35–40.

*These references are the 46 impact evaluations that the Independent Evaluation Group reviewed for this study.

*Iannotti, Lara L., Nelly Zavaleta, Zulema Leon, Anuraj H. Shankar, and Laura E. Caulfield. 2008. "Maternal Zinc Supplementation and Growth in Peruvian Infants." *American Journal of Clinical Nutrition* 88: 154–60.

Independent Evaluation Group (IEG). 2009. *Improving Effectiveness and Outcomes for the Poor in Health, Nutrition, and Population.* Washington, DC: World Bank.

———. 2008. *What Works in Water Supply and Sanitation? Lessons from Impact Evaluation.* Washington, DC: World Bank.

*———. 2005. *Maintaining Momentum to 2015? An Impact Evaluation of Interventions to Improve Maternal and Child Health and Nutrition in Bangladesh.* Washington, DC: World Bank.

Karim, Rezaul, Sascha A. Lamstein, M. Akhtaruzzaman, Khan Maududur Rahman, and Nurul Alam. 2003. *The Bangladesh Integrated Nutrition Project Community-Based Nutrition Component: Endline Evaluation. Final Report.* Institute of Nutrition and Food Science, University of Dhaka, Dhaka, Bangladesh; and Friedman School of Nutrition Science and Policy, International Food and Nutrition Center, Tufts University, Medford, MA.

*Kazianga, Harounan, Damien de Walque, and Harold Alderman. 2009. "Educational and Health Impacts of Two School Feeding Schemes: Evidence from a Randomized Trial in Rural Burkina Faso." Policy Research Working Paper 4976, World Bank, Washington, DC.

King, Elizabeth M., and Jere R. Behrman. 2009. "Timing and Duration of Exposure in Evaluations of Social Programs." *World Bank Research Observer* 24 (1): 55–82.

Kramer, Michael S. 1987. "Determinants of Low Birth Weight: Methodological Assessment and Meta-Analysis." *Bulletin of the World Health Organization* 65 (5): 663–737.

*Leon, Mauricio, and Stephen D. Younger. 2007. "Transfer Payments, Mothers' Income and Child Health in Ecuador." *Journal of Development Studies* 43 (6): 1126–43.

*Leroy, Jeff L., Armando Garcia-Guerra, Raquel Garcia, Clara Dominguez, Juan Rivera, and Lynnette Neufeld. 2008. "The Oportunidades Program Increases Linear Growth of Children Enrolled at Young Ages in Urban Mexico." *The Journal of Nutrition* 138: 793–98.

Levinson, F. J., and J. E. Rohde. 2005. Letter to the Editor. *Health Policy and Planning* 20 (6): 405–06.

*Linnemayr, Sebastian, and Harold Alderman. 2008. "Almost Random: Evaluating a Large-Scale Randomized Nutrition Program in the Presence of Crossover." Policy Research Working Paper 4784, World Bank, Washington, DC.

*Macours, Karen, Norbert R. Schady, and Renos Vakis. 2008. "Cash Transfers, Behavioral Changes and the Cognitive Development of Young Children: Evidence from a Randomized Experiment." Policy Research Working Paper 4759, World Bank, Washington, DC.

*Maluccio, John A., and Rafael Flores. 2005. "Impact Evaluation of a Conditional Cash Transfer Program: The Nicaraguan Red de Protección Social." Research Report 14, International Food Policy Research Institute, Washington, DC.

*Masanja, H., J. Armstrong Schellenberg, D. De Savigny, H. Mshinda, and C. Victora. 2005. "Impact of Integrated Management of Childhood Illness on Inequalities in Child in Child Health in Rural Tanzania." *Health Policy and Planning* 20 (Supplement 1): 77–84.

McGregor, Sally, Yin Bun Cheung, Santiago Cueto, Paul Glewwe, Linda Richter, Barbara Strupp, and the International Child Development Steering Group. 2007. "Development Potential in the First 5 Years for Children in Developing Countries." *The Lancet* 369: 60–70.

*Menéndez, C., A. Bardaji, B. Sigauque, C. Romagosa, S. Sanz, E. Serra-Casas, E. Macete, A. Berenguera, C. David, D. Dobaño, D. Naniche, A. Mayor, J. Ordi, I. Mandomando J. J. Aponte, S. Mabunda, and P. L. Alonso. 2008. "A Randomized Placebo-Controlled Trial of Intermittent Preventive Treatment in Pregnant Women in the Context of Insecticide-Treated Nets Delivered through the Antenatal Clinic." *PLoS ONE*, 3 (4): e1934.

*Miguel, Edward, and Michael Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72 (1): 159–217.

*Morris, S., P. Olinto, R. Flores, E. Nilson, and A. Figueiró. 2004. "Conditional Cash Transfers Are Associated with a Small Reduction in the Rate of Weight Gain of Preschool Children in Northeast Brazil." *The Journal of Nutrition* 134: 2336–41

*Osrin, David, Anjana Vaidya, Yagya Shresta, Ram Bahadur Baniya, Dharma S. Manandhar, Ramesh K. Adhikari, Suzanne Filteau, Andrew Tomkins, and Anthony M. de L. Costello. 2005. "Effects of Antenatal Multiple Micronutrient Supplementation on Birthweight and Gestational Duration in Nepal: Double-Blind, Randomized Controlled Trial." *The Lancet* 365: 955–62.

*Paxson, Christina, and Norbert R. Schady. Forthcoming. "Does Money Matter? The Effects of Cash Transfers on Child Health and Development in Rural Ecuador." *Economic Development and Cultural Change*.

*Penny Mary E., Hilary M. Creed Kanashiro, Rebecca C. Robert, M. Rocio Narro, Laura E. Claufield, and Robert E. Black. 2005. "Effectiveness of an Educational Intervention Delivered through the Health Services to Improve Nutrition in Young Children: A Cluster-Randomized Controlled Trial." *The Lancet* 365: 1863–71.

Pelletier, David, Meera Shekar, L. Du, and Kees Kostermans. 2005. "Bangladesh Integrated Nutrition Project: Effectiveness and Lessons." Bangladesh Development Series, Paper No. 7. World Bank, Dhaka, Bangladesh.

*These references are the 46 impact evaluations that the Independent Evaluation Group reviewed for this study.

*Quisumbing, Agnes R. 2003. "Food Aid and Child Nutrition in Rural Ethiopia." *World Development* 31 (7): 1309–24.

*Ramakrishnan, Usha, Teresa González-Cossío, Lynnette M. Neufeld, Juan Rivera, and Reynaldo Martorell. 2003. "Multiple Micronutrient Supplementation during Pregnancy Does Not Lead to Greater Infant Birth Size Than Does Iron-Only Supplementation: A Randomized Controlled Trial in a Semi-Rural Community in Mexico." *American Journal of Clinical Nutrition* 77: 720–25.

Ramakrishnan, Usha. 2004. "Nutrition and Low Birth Weight: From Research to Practice." *American Journal of Clinical Nutrition* 79: 17–21.

Ravallion, Martin. 2009a. "Should the Randomistas Rule?" *Economists' Voice* 6 (2): 1–5.

———. 2009b. "Evaluation in the Practice of Development." *World Bank Research Observer* 24 (1): 29–54.

———. 2005. "Evaluating Anti-Poverty Programs." Policy Research Working Paper 3625, World Bank, Washington, DC.

*Rivera, Juan A., Daniela Stores-Alvarez, Jean-Pierre Habicht, Teresa Shamah, and Salvador Villalpando. 2004. "Impact of the Mexican Program for Education, Health, and Nutrition (PROGRESA) on Rates of Growth and Anemia in Infants and Young Children: A Randomized Effectiveness Study." *Journal of the American Medical Association* 291 (21): 2563–70.

Ruel, Marie, and John Hoddinott. 2008. "Investing in Early Childhood Nutrition." IFPRI Policy Brief 8. International Food Policy Research Institute, Washington, DC.

*Ruel, Marie, Purnima Menon, Jean-Pierre Habicht, Cornelia Loechl, Gilles Bergeron, Gretel Pelto, Mary Arimond, John Maluccio, Lesly Michaud, and Bekele Hankebo. 2008. "Age-Based Preventive Targeting of Food Assistance and Behavior Change and Communication for Reduction of Childhood Under Nutrition in Haiti: A Cluster Randomized Trial." *The Lancet* 371 (9612): 588–95.

Ruiz, Fernando. 1996. *Estudio de Costos del Proyecto Integral Desarrollo Infantil (PIDI)*. La Paz, Bolivia: Ruiz Guissani Consultores.

Sack, D. A., S. K. Roy, T. Ahmed, and G. Fuchs. 2005. Letter to the Editor. *Health Policy and Planning* 20 (6): 406–07.

Sall, Macky, and Jacques Sylla. 2005. "African Prime Ministers Take Lead in Child Survival." *The Lancet* 366: 1988–89.

*Santos, Iná, Cesar G. Victora, José Martines, Helen Gonçalves, Denise P. Giante, Neiva J. Valle, and Gretel Pelto. 2001. "Nutrition Counseling Increases Weight Gain among Brazilian Children." *Journal of Nutrition* 131: 2866–73.

*Schipani, Steven, Frits van der Haar, Sangsom Sinawat, and Kandavasee Maleevong. 2002. "Dietary Intake and Nutritional Status of Young Children in Families Practicing Mixed Home Gardening in Northeast Thailand." *Food and Nutrition Bulletin* 23(2): 175–180.

*Schroeder, Dirk G., Helena Pachón, Kirk A. Dearden, Tran Thu Ha, Tran Thi Lang, and David R. Marsh. 2002. "An Integrated Child Nutrition Intervention Improved Growth of Younger, More Malnourished Children in Northern Viet Nam." *Food and Nutrition Bulletin* 23 (4) (supplement): 50–58.

Shekar, Meera, on behalf of 17 signatories. 2008. "Delivery sciences in nutrition." Letter to the editor. *The Lancet* 371: 1751.

Shrimpton, Roger, Cesar G. Victora, Mercedes de Onis, Rosangela Cost Lima, Monika Blossner, and Graeme Clugston. 2001. "Worldwide Timing of Growth Faltering: Implications for Nutritional Interventions." *Pediatrics* 107 (5): e75.

Smith, Lisa C., and Lawrence Haddad. 2002. "How Potent Is Economic Growth in Reducing Undernutrition? What Are the Pathways of Impact? New Cross-Country Evidence." *Economic Development and Cultural Change* 51 (1): 55–76.

*Stifel, David, and Harold Alderman. 2006. "The 'Glass of Milk' Subsidy Program and Malnutrition in Peru." *World Bank Economic Review* 20 (3): 421–48.

UNICEF (United Nations Children's Fund). 2007. "Progress for Children: A World Fit for Children Statistical Review." http://www.unicef.org/publications/index_42117.html.

———. 1990. "Strategy for Improved Nutrition of Children and Women in Developing Countries." UNICEF Policy Review, New York.

UNICEF (United Nations Children's Fund) and World Health Organization. 2004. *Low Birthweight: Country, Regional and Global Estimates.* UNICEF: New York.

Van de Poel, Ellen, Ahmad Reza Hosseinpoor, Niko Speybroeck, Tom Van Ourti, and Jeannette Vega. 2008. "Socioeconomic Inequality in Malnutrition in Developing Countries." *Bulletin of the World Health Organization* 86 (4): 282–91.

Victora, Cesar, Jean-Pierre Habicht, and Jennifer Bryce. 2004. "Evidence-Based Public Health: Moving Beyond Randomized Trials." *American Journal of Public Health* 94 (3): 400–05.

Walker, Susan P., Theodore D. Wachs, Julie Meeks Gardner, Betsy Lozoff, Gail A. Wasserman, Ernesto Pollitt, Julie A. Carter, and the International Child Development Steering Group. 2007. "Child Development: Risk Factors for Adverse Outcomes in Developing Countries." *The Lancet* 369: 145–57.

*Waters, Hugh R., Mary E. Penny, Hilary M. Creed-Kanashiro, Rebecca C. Robert, Pocio Naro, Jeff Willis, Laura E. Caulfield, and Robert E. Black. 2006. "The Cost-

*These references are the 46 impact evaluations that the Independent Evaluation Group reviewed for this study.

Effectiveness of a Child Nutrition Education Programme in Peru." *Health Policy and Planning* 21: 257–64.

West, Stephen G., Naihua Duan, Willo Pequegnat, Paul Gaist, Don C. Des Jarlais, David Holdgrave, Jose Szapocznik, Martin Fishbein, Bruce Rapkin, Michael Clatts, and Patricia Dolan Mullen. 2008. "Alternatives to the Randomized Controlled Trial." *American Journal of Public Health* 98 (8): 1359–66.

*White, Howard, and Edoardo Masset. 2007. "Assessing Interventions to Improve Child Nutrition: A Theory-Based Impact Evaluation of the Bangladesh Integrated Nutrition Project." *Journal of International Development* 19 (5): 627–52.

White, Howard. 2005. "Comment on contributions regarding the impact of the Bangladesh Integrated Nutrition Project." *Health Policy and Planning* 20*:* 408–410.

WHO (World Health Organization). 1995. "Physical Status: The Use and Interpretation of Anthropometry." Report of a WHO Expert Committee, Technical Report Series 854, Geneva.

World Bank. 2009a. *Global Monitoring Report 2009: A Development Emergency.* Washington, DC: World Bank.

———. 2009b. "Implementation of the World Bank's Strategy for Health, Nutrition, and Population (HNP) Results: Achievements, Challenges, and the Way Forward—Progress Report." Health, Nutrition, and Population, Human Development, World Bank, Washington, DC.

———. 2009c. "Methodologies to Evaluate the Impact of Large-Scale Nutrition Projects." *Doing Impact Evaluation,* No. 13. World Bank, Washington, DC.

———. 2008. "Project Appraisal Document on a Proposed Loan in the amount of US$636.5 million to the Republic of Colombia for a Support for the Second Phase of the Expansion of the Program of Conditional Transfers-*Familias en Accion* Project." Report No. 45377-CO, World Bank, Washington, DC.

———. 2007a. *Healthy Development: The World Bank Strategy for Health, Nutrition, and Population Results.* Washington, DC: World Bank.

———. 2007b. "Implementation Completion and Results Report on a Credit in the Amount of SDR 11.8 Million (US$14.7 million equivalent) to the Republic of Senegal in Support of the First Phase Nutrition Enhancement Program." Report No. ICR000107, World Bank, Washington, DC.

———. 2007c. "Impact Evaluation for Microfinance." *Doing Impact Evaluation,* No. 7. Washington, DC: World Bank.

———. 2006a. *Repositioning Nutrition as Central to Development: A Strategy for Large-Scale Action.* Washington, DC: World Bank.

———. 2006b. "Implementation Completion Report (FSLT-70500) on a Loan in the Amount of US$150.0 Million to the Republic of Colombia for a Human Capital Protection Project." Report No. 35614, World Bank, Washington, DC.

———. 2005a. "The Bangladesh Integrated Nutrition Project: Effectiveness and Lessons." Bangladesh Development Series*,* Paper No. 8, Dhaka, Bangladesh.

———. 2005b. "Project Appraisal Document on a Proposed Loan in the Amount of US$86.4 Million to the Republic of Colombia for a Social Safety Net Project." Report No. 31880-CO, World Bank, Washington, DC.

———. 2004. "Implementation Completion Report (IDA-25310) on a Credit in the Amount of SDR 35.8 Million (US$50.7 million equivalent) to the Republic of Bolivia for an Integrated Child Development Project." Report No. 30209, World Bank, Washington, DC.

———. 2003. "Program Document for a Proposed Programmatic Human Development Reform Loan I in the Amount of US$50.00 Million to the Republic of Ecuador." Report No. 25791-EC, World Bank, Washington, DC.

———. 2001. "Project Appraisal Document on a Proposed Loan in the Amount of US$150.0 Million to the Republic of Colombia for a Human Capital Protection Project." Report No. 21608-CO, World Bank, Washington, DC.

———. 1998. "Project Appraisal Document on a Proposed Credit in the Amount of SDR 20.4 Million Equivalent to the Republic of Madagascar for a Community Nutrition II Project." Report No. 17507, World Bank, Washington, DC.

———. 1993. "Staff Appraisal Report: Bolivia Integrated Child Development Project." Report No. 11905-BO, World Bank, Washington, DC.

*Yamano, Takashi, Harold Alderman, and Luc Christiaensen. 2005. "Child Growth, Shocks, and Food Aid in Rural Ethiopia." *American Journal of Agricultural Economics* 87 (2): 273–88.

*Zeng, L., Y. Cheng, S. Dang, H. Yan, M. J. Dibley, S. Chang, and L. Kong. 2008. "Impact of Micronutrient Supplementation during Pregnancy on Birthweight, Duration of Gestation, and Perinatal Mortality in Rural Western China: Double-Blind Cluster Randomized Controlled Trial." *British Medical Journal* 337: a2001.

*These references are the 46 impact evaluations that the Independent Evaluation Group reviewed for this study.

IEG Publications

Annual Review of Development Effectiveness 2009: Achieving Sustainable Development
Addressing the Challenges of Globalization: An Independent Evaluation of the World Bank's Approach to Global Programs
Assessing World Bank Support for Trade, 1987–2004: An IEG Evaluation
Books, Building, and Learning Outcomes: An Impact Evaluation of World Bank Support to Basic Education in Ghana
Bridging Troubled Waters: Assessing the World Bank Water Resources Strategy
Climate Change and the World Bank Group—Phase I: An Evaluation of World Bank Win-Win energy Policy Reforms
Debt Relief for the Poorest: An Evaluation Update of the HIPC Initiative
A Decade of Action in Transport: An Evaluation of World Bank Assistance to the Transport Sector, 1995–2005
The Development Potential of Regional Programs: An Evaluation of World Bank Support of Multicountry Operations
Development Results in Middle-Income Countries: An Evaluation of World Bank Support
Doing Business: An Independent Evaluation—Taking the Measure of the World Bank–IFC Doing Business Indicators
Egypt: Positive Results from Knowledge Sharing and Modest Lending—An IEG Country Assistance Evaluation 1999–2007
Engaging with Fragile States: An IEG Review of World Bank Support to Low-Income Countries Under Stress
Environmental Sustainability: An Evaluation of World Bank Group Support
Evaluation of World Bank Assistance to Pacific Member Countries, 1992–2002
Extractive Industries and Sustainable Development: An Evaluation of World Bank Group Experience
Financial Sector Assessment Program: IEG Review of the Joint World Bank and IMF Initiative
From Schooling Access to Learning Outcomes: An Unfinished Agenda—An Evaluation of World Bank Support to Primary Education
Hazards of Nature, Risks to Development: An IEG Evaluation of World Bank Assistance for Natural Disasters
How to Build M&E Systems to Support Better Government
IEG Review of World Bank Assistance for Financial Sector Reform
An Impact Evaluation of India's Second and Third Andhra Pradesh Irrigation Projects:
A Case of Poverty Reduction with Low Economic Returns
Improving Effectiveness and Outcomes for the Poor in Health, Nutrition, and Population
Improving the Lives of the Poor through Investment in Cities
Improving Municipal Management for Cities to Succeed: An IEG Special Study
Improving the World Bank's Development Assistance: What Does Evaluation Show:
Maintaining Momentum to 2015: An Impact Evaluation of Interventions to Improve
Maternal and Child Health and Nutrition Outcomes in Bangladesh
New Renewable Energy: A Review of the World Bank's Assistance
Pakistan: An Evaluation of the World Bank's Assistance
Pension Reform and the Development of Pension Systems: An Evaluation of World Bank Assistance
The Poverty Reduction Strategy Initiative: An Independent Evaluation of the World Bank's Support Through 2003
The Poverty Reduction Strategy Initiative: Findings from 10 Country Case Studies of World Bank and IMF Support
Power for Development: A Review of the World Bank Group's Experience with Private Participation in the Electricity Sector
Public Sector Reform: What Works and Why? An IEG Evaluation of World Bank Support
Small States: Making the Most of Development Assistance—A Synthesis of World Bank Findings
Social Funds: Assessing Effectiveness
Sourcebook for Evaluating Global and Regional Partnership Programs
Using Knowledge to Improve Development Effectiveness: An Evaluation of World Bank
Economic and Sector Work and Technical Assistance, 2000–2006
Using Training to Build Capacity for Development: An Evaluation of the World Bank's Project-Based and WBI Training
The Welfare Impact of Rural Electrification: A Reassessment of the Costs and Benefits—An IEG Impact Evaluation
World Bank Assistance to Agriculture in Sub-Saharan Africa: An IEG Review
World Bank Assistance to the Financial Sector: A Synthesis of IEG Evaluations
World Bank Group Guarantee Instruments 1990–2007: An Independent Evaluation
The World Bank in Turkey: 1993–2004—An IEG Country Assistance Evaluation
World Bank Engagement at the State Level: The Cases of Brazil, India, Nigeria, and Russia

All IEG evaluations are available, in whole or in part, in languages other than English.
For our multilingual section, please visit http://www.worldbank.org/ieg.

ISBN 978-0-8213-8406-0

SKU 18406

IEG WORLD BANK | IFC | MIGA
Independent Evaluation Group