Assessing Poverty Impacts of Agricultural Research: Methods and Challenges for CGIAR

CGIAR

Independent Science and Partnership Council

Standing Panel on Impact Assessment (SPIA)

ABOUT ISPC

The Independent Science and Partnership Council (ISPC) is a part of CGIAR, a global research partnership for a food secure future dedicated to achieving a world free of poverty, hunger and environmental degradation. The ISPC provides independent advice to CGIAR System Council which is comprised of funders and representatives of developing countries. The mission of the ISPC is to help strengthen the quality, relevance, and impact of CGIAR research by enhancing the System Council's capacity to make evidence-based decisions in support of effective agricultural research programs for sustainable development.

ABOUT SPIA

The Standing Panel on Impact Assessment (SPIA) is a sub-group of the ISPC. SPIA's mandate is to provide CGIAR with timely, objective, and credible information on the impacts of research at the system level; provide support to and complement CGIAR centers in their *ex post* impact assessment activities; and, provide feedback to planning, monitoring and evaluation functions in CGIAR.

Independent Science and Partnership Council (ISPC)

http://ispc.cgiar.org

Cover image: ©FAO/Simon Maina

Design and layout: Macaroni Bros

Editor: Heidi Fritschel

Douglas Gollin
Lilli Teresa Probst
Elizabeth Brower

Assessing Poverty Impacts of Agricultural Research: Methods and Challenges for CGIAR

June 2018

TABLE OF CONTENTS

EXECUTIVE SUMMARY

For more than 50 years, the international community has viewed agricultural research as a key instrument in the quest to improve livelihoods in the developing world. Governments (both directly and through multilateral institutions) and philanthropies have poured hundreds of millions of dollars into publicly funded agricultural science targeted to the problems of the developing world. Much of this publicly-funded agricultural science has taken place in CGIAR and its component centers and research programs. The work done by CGIAR and its partners has been hugely important; probably the best measure of its impact is the sheer extent of the adoption of CGIAR technologies by farmers and by governments. However, questions remain on how much CGIAR's research has actually reduced poverty, and how best to describe, measure and assess such impacts.

This document analyzes and discusses different research methods that can contribute to an understanding of poverty impacts attributable to agricultural research. It takes it as a central premise that agricultural research can and does contribute to changes in poverty, and that there are numerous pathways through which such impacts are generated. The purpose of this study is to provide guidance for researchers who intend to carry out poverty impact assessments. It reviews methods used and lessons from a number of past impact assessments—including some commissioned or carried out by the Standing Panel on Impact Assessment (SPIA)—that have attempted to document poverty impacts. This document might also be a useful resource for research managers, including donors, who want to know what they can realistically say about the poverty impacts of their investments.

Our view is that there is a long and complex causal chain leading from agricultural research to poverty impacts. To understand this causal chain, we will need to employ many methods—from randomized controlled trials to descriptive histories. Different methods can provide information at different spatial and temporal scales. There are roles for both quantitative and qualitative methods; good research of any kind can add to our understanding. Researchers will need to choose appropriate research designs for the setting, the stage of adoption, and the scale of impact, among other factors.

FOREWORD

How can we know if our investments in agricultural research are effective in reducing poverty? This is a frequent question donors pose to the ISPC and CGIAR system, and in an era of tightening budgets it has become even more critical to answer. This paper responds to that question by providing an overview of approaches to agricultural research impact assessment, focusing mostly on research that leads to the development of new crop varieties but with implications for the broader research agenda. In a nutshell, it tells us that there are many methods for assessing the impacts of agricultural research on poverty levels, and that they have their pros and cons depending on specific circumstances. More important than the choice of method, a good impact assessment must be built upon a clear understanding of how the research outputs are expected to affect poverty and its various dimensions—e.g., the impact pathway. These pathways are generally complicated and multi-layered, and unfold over time—often in unexpected ways. Thus, questioning widespread assumptions about how agricultural research can contribute to poverty reduction is essential for designing robust assessments of its impact.

The paper provides a comprehensive overview of how to conduct robust assessments of the impact of agricultural research on reducing poverty—and, also gives insights and guidance on better designing agricultural research to contribute to poverty reduction. The main research production paradigm adopted in the paper is the classic top-down technology design and dissemination model that is still relevant for some types of agricultural research, but will not be applicable in other important areas of agricultural research. For example, assessing impacts of policy-oriented research, or programs where new knowledge and innovations are co-produced by scientists and farmers in an action research mode may require different approaches and methods than the ones covered in this paper. Nonetheless, this paper is relevant for a range of audiences: from agricultural researchers engaged in new technology development, as well as more broadly in the field of agricultural development and poverty reduction, to impact assessment specialists, and donors who are concerned with realizing poverty impacts from investments in agricultural research. Different segments of the paper are more relevant for each of these audiences, although the pragmatic and comprehensible style throughout the paper makes even the more specialized topics accessible to all.

Section 2 starts by outlining major concepts in the relationship between agricultural research and the technologies it generates, and the various ways in which these may, or may not reduce poverty. The bottom line in this section is that there are several ways in which agricultural research and technologies might contribute to, or constrain, efforts to reduce poverty. It is important to think through the multiple potential pathways of impact before designing an impact assessment. The authors start by questioning widely held assumptions about agricultural technologies and their effect on the wealth-generating capacity of poor people, focussing on yield-enhancing improved crop varieties. Often, these are assumed to be a primary means of increasing poor farmers' incomes. The authors walk the reader through examples where this assumption does not hold; where there are no poverty impacts of a yield increase—or even negative impacts. Likewise, the authors examine another common assumption that underlies the design of agricultural research programs—that yield increases will reduce food prices and thus benefit poor consumers. Here again, however, the paper points out situations where the expected effect may not materialize, and instead perverse effects could come about. This analysis is nicely summarized in Table 1, where the likely effects of productivity-enhancing innovations on poverty are traced across a range of different situations for producers, consumers and agricultural labourers. The analysis then goes beyond research innovations that increase yields to consider a variety of different potential effects and how they may affect poverty.

Section 3 lays out the key methodological issues in designing a robust assessment of poverty impacts from agricultural research. Getting into the nitty gritty of impact assessment methods, this section is most relevant to impact assessment specialists—be they academic or practitioners. It also includes a review of relevant, recent literature on the impact of agricultural research and technologies on poverty. There are several empirical considerations for achieving a robust impact assessment from sample design to statistical approach. The authors give great emphasis to making sure that impact assessment actually samples the right population—e.g., the people whose lives the technology could potentially affect. Methods for discerning whether poverty impacts are due to the adoption of the technology or some other underlying characteristic of the population are presented, as well as those for attribution of impacts. The methods described fall into three broad categories: micro level observational and experimental studies; studies looking at broader development outcomes, including indirect impacts; and macro-level studies that describe the effect of growth in agricultural productivity on poverty. Here too, the pros and cons of each from a methodological point of view are discussed. The section includes some discussion of qualitative methods, although the emphasis is clearly on the quantitative approaches. While this discussion will be of most interest to impact assessment specialists, the summary of the results from literature review in Appendix 1 has much broader appeal to agricultural development practitioners and donors. In particular, it indicates gaps in evidence on specific agricultural research to poverty pathways and related outcomes.

In the following section, the authors take on the larger issue of research design—which too often is not given adequate consideration in research methodology deliberations. The authors discuss how best to identify the potential geographical area and populations of interest, the rates of uptake, the timeframe for expected outcomes and impacts, the expected diffusion pathways, and the potential heterogeneity of impacts. All of these factors, which essentially constitute the expected impact pathway, should be agreed upon before any technology is introduced to farmers. The objective of good research design is to be as forward looking as feasible, even as impact assessments looks backward. The authors call for a joint effort between scientists involved in the technology development and social scientists involved in the impact assessment in developing the research design—which all too often is not the case.

One of the most interesting aspects of the paper is the discussion across various sections on the use of randomization, which is a much discussed subject in the impact assessment world. The key issue for designing a robust impact assessment is to have a clear idea of the counterfactual: i.e., what would have happened if the innovation was not introduced. Randomization is a good way to build a suitable counterfactual, but not always feasible or even the best way to create a counterfactual. The authors suggest other possible approaches for varying circumstances.

Overall, the paper makes the case that assessing poverty impacts of agricultural research and the technologies it produces is feasible, but is also complicated, requiring a long-term commitment and sufficient resources to be credible. As the premium global public sector agricultural research system with poverty reduction as a key objective, CGIAR has an important role to play in showing how its research can and does contribute to poverty reduction, but also how a robust and effective impact assessment approach can play an integral role in achieving an effective research system.

Leslie Lipper
Executive Director, ISPC Secretariat

1 INTRODUCTION

For more than 50 years, the international community has viewed agricultural research as a key instrument in the quest to improve livelihoods in the developing world. Governments (both directly and through multilateral institutions) and philanthropies have poured hundreds of millions of dollars, at today's prices, into publicly funded agricultural science targeted to the problems of the developing world.

Much of this publicly-funded agricultural science has taken place in CGIAR and its component centers and research programs. This network of research institutions, along with numerous partners, has worked on a broad range of topics. The most prominent area of research has been genetic improvement in crops, but scientists in CGIAR have also worked on such disparate topics as livestock breeding, vaccine development, and animal management; natural resource management; water and irrigation; agronomic practices; the development of appropriate agricultural machines; food and agriculture policy; and a host of other issues. The work done by CGIAR and its partners has been hugely important; probably the best measure of its impact is the sheer extent of the adoption of CGIAR technologies by farmers and by governments.

But how much has CGIAR's research actually served to reduce poverty? In an era of competing claims on donor resources—and in an era when there is intense demand for evidence-based policymaking—how can we measure the impact of agricultural research on poverty? This document analyzes and discusses some of the different research methods that can contribute to an understanding of poverty impacts attributable to agricultural research. The purpose of this paper is to provide guidance for researchers who intend to carry out poverty impact assessments. We also imagine that this document might be a useful resource for research managers, including donors, who want the option, in the future, to assess the poverty impacts of investments they are making today.

This paper takes it as a central premise that agricultural research can and does contribute to changes in poverty. There are numerous pathways through which agricultural research can generate impacts on poverty. We do not aim to summarize these theoretical pathways; neither do we intend to review the evidence for the magnitude of these effects. Instead, the aim of this paper is to discuss methodological approaches to describing, measuring, and assessing poverty impacts.

This paper is intended neither as an econometrics text that might review specific techniques nor as a manifesto that will insist on the correctness of any particular approach

Our view is that there is a long and complex causal chain leading from agricultural research to poverty impacts. To understand this causal chain, we will need to employ many methods—from randomized controlled trials to descriptive histories. Different methods can provide information at different spatial and temporal scales. There are roles for both quantitative and qualitative methods; good research of any kind can add to our understanding.

Our primary focus in this document is on research design. What kinds of data must be collected to measure poverty impacts? How should these data be analyzed? At what stage in the development of a new technology, or at what stage of its diffusion to farmers, should surveys be carried out? What are the pitfalls and challenges of different methodological approaches? What are appropriate methods for different questions and different settings? How do the costs of data collection and analysis affect the choice of research methods? The answers to these questions will surely vary from context to context. No single correct formula will apply in every setting. Instead, researchers will need to choose appropriate research designs for the setting, the stage of adoption, and the scale of impact, among other factors.

In the pages that follow, we review a number of specific research methodologies—both quantitative and qualitative—that can offer insights into poverty impacts. Depending on the setting, some approaches may be more promising than others for observing and assessing poverty impacts. We examine previous literature that uses these different methods, and we try to understand how the appropriate research design relates to the spatial and temporal scale at which poverty impacts are sought. Research design may also depend on the type of technology and its stage of diffusion to farmers.

The structure of this paper is as follows. In section 2, we trace out a conceptual framework that incorporates many of the pathways through which agricultural research can lead to poverty impacts. These pathways define the research agenda; if we are to assess poverty impacts, we must know where to look and what potential causal relationships to explore. Section 3 then reviews a number of past studies—including some commissioned or carried out by the Standing Panel on Impact Assessment (SPIA)—that have attempted to document poverty impacts. This is not a comprehensive literature review, and we do not focus on the findings of these studies. Instead, we point out the methods used in past studies and some of the lessons learned for future research design.

Section 4 talks about the broader issue of research design. We argue that statistical methods are less of a challenge for impact assessment than careful research design is. The challenge is to find internal consistency between the research question, the chosen methods, and the data that are collected. A sensible research design begins with an effort to establish appropriate counterfactuals and to set up data collection such that poverty impacts can eventually be measured.

We take up the issue of research design again in the conclusions, which are presented in section 5. The report has three central findings:

First, poverty impact assessment is difficult, and the sheer complexity of the causal connections between agricultural research and poverty levels makes it exceedingly difficult to find clear and direct evidence of these causal links. The best we can do in most cases is to provide a portfolio of evidence, drawn from multiple sources and methods, which can be synthesized into a reasonably complete overall picture.

Second, to produce a compelling analysis of poverty impacts, the research effort must begin at early stages of the introduction of a new technology, and it may need to last for many years. We argue that the key issue for effective assessment of poverty impacts is research design. Without thoughtful research design at the early stages of technology introduction, there is no econometric technique that can provide convincing evidence after the fact.

Third, because poverty impact assessment must begin with a coherent research design, it requires close cooperation between social scientists and the developers of the underlying technologies—in this case, CGIAR scientists and managers. This cooperation should cover tasks such as identifying potential areas of impact, collecting baseline data, and, especially, carefully measuring on-farm productivity benefits under accurately representative conditions.

Finally, although this paper does not take up the issue, it is vital that information about poverty impacts feed back into the research-planning and priority-setting process. The discussion hitherto has tended to assume, on the basis of little evidence, that the poverty impacts of agricultural research come through two main channels: (1) the tendency of productivity gains to drive down food prices, helping poor consumers; and (2) the income gains accruing to poor farmers from reductions in the unit cost of production (or equivalently from increasing output for the same inputs). Both of these channels imply that the main goal of research, from a poverty reduction perspective, should be to increase productivity, measured in quantity terms. This paper suggests, however, that other channels may be important for poverty reduction, such as those operating through labor markets. The complex causal relationships between research and poverty reduction may imply a more complicated set of research priorities, with greater need for case-by-case analysis of the poverty impacts of research. In the next section, we explore these causal relationships in greater detail.

2 CONCEPTUAL FRAMEWORK

One of CGIAR's three system-level objectives (SLOs) is to use agricultural science to reduce poverty in developing countries. This goal reflects growing pressure in the international development community for organizations to hold themselves accountable for their effectiveness. Institutions like CGIAR are expected to define objectives such that their progress can be monitored transparently and their impact on development can be made clear.

There are many theoretical reasons to believe that agricultural research can contribute to poverty reduction. Moreover, there are many plausible pathways through which this impact can take place. But documenting and quantifying these impacts is not easy. The timescales on which agricultural research has its impacts are generally long, and poverty reduction involves highly complex social and economic processes. Poverty has deep roots, and poverty reduction depends on many factors other than agricultural research.

Finding causal connections between agricultural research and poverty outcomes is difficult, in large part because of this long causal chain, which we illustrate in Figure 1 (suggested by Robert W. Herdt). Research often begins in laboratories and on agricultural experiment stations—though increasingly it also draws on participatory approaches and farmer knowledge. New technologies and policy or institutional innovations are tried and tested, modified, and evaluated. At some stage, they are disseminated to farmers or other agricultural sector actors, who may then undertake their own informal experimentation before adopting. Even then, farmers may try out new technologies on a limited basis for years, or even decades, before fully adopting. Even at that point, the impact on producer livelihoods and well-being is complex. New production technologies have immediate (first-round) effects on the resources used in the production system and perhaps on consumers, but may also prompt far-reaching rearrangements of household livelihood strategies, affecting patterns of work, sources of income, and even decisions about education and migration. It may take many years before the net effects on household well-being and poverty are realized.

The timeline from research to poverty impact may thus span several decades. Over this length of time, it is difficult to distinguish the effects of agricultural research from a host of other forces that are also at work. Policy changes, infrastructure projects, education and health interventions, and any number of other factors will have changed during the time period, making it extremely difficult to attribute any reductions in poverty to the effects of agricultural research.

The problem is not insurmountable. Well-chosen methods can help identify some of the specific links in the causal chain. Good qualitative and descriptive work can provide useful insights into the broader social and historical forces that have accompanied technology adoption and diffusion. The tools of social science can help us understand both the broad patterns of poverty impact and the narrower outcomes of specific episodes of technology change.

2.1 PATHWAYS AND CHANNELS TO RESEARCH IMPACT: PRODUCTIVITY CHANGE AND BEYOND

Agricultural research can affect poverty through a number of plausible pathways. The main pathway is through increases in food productivity, which should, all else being equal, lead to relatively cheaper and more abundant food. This is likely to be good for poor people, who spend large proportions of their income on food and who are the most vulnerable when food is scarce. At the global level, and over long time spans, it is surely better for the poor to be in a world with abundant and inexpensive food than in one where food is scarce.

Although that claim almost certainly holds at the aggregate level, it does not necessarily follow that increases in agricultural productivity are good for all poor individuals. On the contrary, many specific improvements in agricultural productivity are likely to lead some poor people to be worse off. And over shorter time periods, or at the more localized level of a particular community, country, or region, it is not immediately obvious that increases in agricultural productivity will make the poor better off.

To see this, consider a small open economy (an economist's term for a country where prices are set by world markets) where land is heavily concentrated in large farms, and the poor are primarily engaged in wage labor. Suppose that a new technology increases productivity but reduces the demand for labor; this might happen, for example, with certain kinds of mechanical or chemical technologies that economists term as having a labor-saving bias. In this setting, we would expect the productivity improvement to leave the poor worse off, since the decrease in labor demand will lead to a decline in the wages of the poor without any corresponding reduction in the price of food or their cost of living. Technologies that increase yield (kilograms/hectare) are often said to have a land-saving bias, and other categories are recognized.

Thus, the specific poverty impacts caused by a new technology will depend on a complex set of circumstances: there is no easy generalization that applies in all cases. Instead, it is useful to think about the different pathways that may pertain in a given context.

In most cases, our mental model of agricultural research impacts on poverty assumes that research leads to **productivity changes**, measured in terms of changes in the quality or quantity of output that can be obtained from a given bundle of inputs. This is the kind of impact that we might expect to see with a high-yielding crop variety, for instance. The same characterization might also apply to a variety that yielded less but displayed resistance to a particular biotic stress, so that it required lower levels of chemical input or labor. A productivity change might also come from an improved management practice or other shift in a farming system. Economists would use the same framework to think about a change in the quality of output that resulted in an increase in the value of what farmers produce—for example, an improvement in grain quality that leaves yield unaltered. In an economic sense, a change in the unit value of output is essentially equivalent to a change in physical productivity.

But not all research leads to productivity improvements, measured in this way. Other research might lead to changes in the **quality of the resource base** or in the **healthfulness of output** (e.g., aflatoxin control or nutrient fortification). Research of this kind may also have poverty impacts. For instance, the current effort to produce iron-enriched bean varieties might have a poverty-reducing impact if it succeeds in alleviating the burdens of iron deficiency and anemia that together limit the labor supply of poor people. One of the hopes for Golden Rice is that it can reduce the prevalence of blindness; this is obviously worthwhile in itself, but it would also be expected to have an additional benefit in reducing the poverty and dependence of those afflicted by blindness and the families that support them.

Other types of research impacts that potentially impact poverty—but do not involve productivity increases in the narrow sense—include technologies that **reduce the time burdens** of production and processing, thereby freeing labor for other employment activities, and research that **improves policies and institutions**.

2.2 POPULATION GROUPS AND CATEGORIES

To understand the impacts of a new agricultural technology on poverty, one might want to understand how it affects each major category of individuals/households. The most obvious categories to consider are consumers, producers, agricultural workers, nonagricultural workers, and landowners. Other categories might also be important in specific cases—for instance, people employed in processing industries. Moreover, the effects on each group will differ depending on whether the economy in question is fully closed to international trade, fully open, or somewhere in between.

The categories are useful from a theoretical sense, but they do not always map neatly onto reality. Many households—and even individuals—fall into several of these categories. Thus, a single individual could easily be a producer of some agricultural goods and a consumer of others; she might own some land but also work off-farm. Such patterns are common and reflect the incentives for diversification of income sources and the flexibility of consumption choices.

It is also important to figure out how the poor are allocated across the different groups. In some economies, the poor may be heavily concentrated in rural areas; they may be landowners and agricultural workers. In other places, the poor may be primarily urban residents, perhaps the self-employed or those working in the informal sector. A new technology will have differing impacts on the poor depending on whether the poor are primarily farmers, urban consumers, or landless agricultural workers. The effects of a technological innovation on poverty overall will thus depend on the distribution of poverty across these population groups. To complicate matters further, the average effects on each group also conceal substantial within-group differences. The categories of "producers" and "consumers" are obviously highly heterogeneous. Landowners may include both the landlords of large estates and those individuals with small holdings. In short, we care about the distribution of technology impacts across these groups but also within each group.

2.3 PATHWAYS TO POVERTY IMPACT

Because of the multiplicity of channels from research to impact, and because of the different population groups involved, there are many possible pathways from agricultural research to poverty impact. This section sketches out some of the key pathways.

2.3.1 Impacts on producers: price-quantity effects

A productivity increase originating from agricultural research has the potential to induce important price and quantity effects on producers. But the nature of these effects will vary substantially depending on the setting: the openness of the economy, the type of technological change (bias), and the rate of adoption, as outlined in Table 1. The biggest beneficial effect will come when poor producers receive a productivity gain from adopting with no corresponding decline in price; that is the special case of a small and fully open economy. In this setting, prices are entirely determined by world markets, so a productivity increase translates entirely into gains in the value of output. (There may, of course, be increased input costs as well.)

The worst case here is that of a poor producer in a small and fully closed economy who does not get the productivity gain (e.g., does not adopt) but does get the price decline. This would be the case, for instance, in a world of differential access to the technology by poor farmers. If rich farmers or those with large landholdings are able to access improved technologies, but poor farmers do not, and if prices drop in response to the increased output from large farms, then the poor farmers can end up absolutely worse off. Even at a global level—the case of a large open economy—similar effects are possible. Evenson and Gollin (2003) argued that something similar happened in the wake of the Green Revolution in global cereal markets, with absolute losses accruing through the price mechanism for farmers who did not adopt or were in areas that did not substantially benefit from the new technologies.

There are of course many intermediate cases between the two extremes described here. In these intermediate cases, producers would receive some mix of productivity gains and price declines. The net effects of productivity increases on poverty, through this price-quantity mechanism, will depend on the degree of openness of the local economy as well as on the distribution of productivity gains across producers of different income levels. The poverty impacts will also be expected to differ depending on the scale of adoption, diffusion, and take-up, along with the actual physical productivity effects of the technology.

Table 1. Nature of the effects of productivity-enhancing innovations

	LIKELY EFFECT OF PRODUCTIVITY-ENHANCING INNOVATION[a]					
			ON AGRICULTURAL LABORERS IF TECHNOLOGY BIAS IS:		ON INCOME OF PRODUCERS WHO ARE:	
ECONOMY TYPE	ON PRODUCT PRICE	ON CONSUMERS	LABOR SAVING	LABOR INTENSIFYING	ADOPTERS	NON-ADOPTERS
Small open	No change	More product, same price => Gains	Less work, lower incomes => Losses	More work, higher income => Gains	Lower unit cost => Gains	No change
Small closed	Reduction	More product, lower price => Gains	Less work, lower wage, less income => Losses	More work lower/higher wage => Uncertain	Lower unit cost Lower/Higher price => Uncertain	Lower sale price => Losses
Large open	Small reduction possible	More product => Gains	Less work Lower wage Less income => Losses	More work, lower/higher wage => Uncertain	Lower unit cost, lower/higher price => Uncertain	Lower sale price => Losses
Large closed	Reduction	More product, lower price => Gains	Less work, lower wage, less income => Losses	More work, lower/higher wage => Uncertain	Lower unit cost, lower/higher price => Uncertain	Lower sale price => Losses

[a] Actual effects depend on market conditions.

The openness of the economy is not entirely an issue of policy choices. It may also reflect the availability of substitutes for the crop in question and the size of the market. For instance, an increase in the productivity of teff in Ethiopia is likely to push down teff prices in Ethiopia, but an increase in Ethiopian maize productivity is likely to have far less impact. The difference between the two cases is that maize is far more extensively traded on global markets, making Ethiopia a small open economy relative to the world maize market, whereas it is probably a large economy with respect to the world teff market.

2.3.2 Impacts on producers: effects other than price-quantity

Some of the poverty impacts from agricultural research on producers may originate through channels that have little to do with the physical quantities of outputs or inputs. As mentioned, a key channel may be changes in the quantity and timing of labor used in farming. Technology changes may free labor for other market work. Poverty impacts may result from the development of appropriate tools and mechanization that reduce drudgery. Poverty may also be affected by innovations that lead to improvements in the physical health of the producers—e.g., by reducing their exposure to chemicals or to risk of injury.

2.3.3 Impacts on consumers: price effects

Perhaps the most obvious poverty impact of agricultural research on consumers occurs through the channel of price and quantity effects. As in the case of price-quantity effects on producers, the importance of this pathway for consumers depends on the openness of the economy and the distance of consumers from the production source. In general, productivity gains will reduce prices and thus benefit consumers in a particular country. Since poor consumers spend disproportionately large fractions of their income on food purchases, a decrease in food prices would normally be expected to benefit the poor.

There are, however, some exceptions to this pattern. In a small, fully open economy, prices are determined by world markets, so we would expect that a productivity increase would not lead to any change in price or poverty. More perverse effects are also possible. For instance, productivity gains may lead to changes in marketing patterns; local producers may begin to sell into regional, national, or international markets, driving up prices for local consumers. Several examples of dairy development have resulted in milk being shipped out of a local catchment to urban areas, with milk prices rising for local consumers as a consequence.

2.3.4 Impacts on consumers: non-price effects

Agricultural research may also affect poverty among consumers in ways other than the sheer quantity of goods produced. For instance, research may reduce poverty through channels such as the quality and characteristics of agricultural goods. For example, nutrient-fortified crops or those that are free from toxins can affect poverty by improving the health of consumers, reducing the amount of money spent on drugs and health care, or reducing days lost to ill health. Other quality characteristics, such as cooking and processing time, may also affect the time required for household labor (primarily of women), allowing greater engagement in other activities that might increase household income and reduce poverty.

2.3.5 Impacts on agricultural workers: wage effects

Many poor people earn their living primarily from agricultural labor—whether on their own farms or on other farms. In general, we would expect agricultural research to affect agricultural workers through the demand for agricultural labor, which in turn affects wages. But technologies may also have indirect effects through the supply of agricultural labor—for instance, by changing households' decisions about how much labor to supply to market activities compared with home production or schooling.

The effects of technology on demand for labor are entirely ambiguous. Productivity changes may increase or decrease the demand for labor—it depends on the technology and the context. It is assuredly not always the case that increases in productivity drive up the demand for labor; in some cases, they reduce the demand for agricultural labor. Wage effects may be complicated by the fact that technologies may alter the demand for specific types of labor, as well as the aggregate demand for labor. For instance, some technologies may increase the demand for family labor but decrease the demand for hired labor (e.g., technologies that involve mechanization) or the opposite. For example, improvements in horticultural technologies may increase the demand for hired labor in picking/harvesting. Some other technologies may alter the demand for the labor of men, women, or children; these will have potentially quite different effects on poverty.

Wage effects may be compounded by bargaining-power effects operating within families or households. Does a decrease in the demand for women's farm labor lead to an increase in women supplying labor to the rural nonfarm sector—potentially reducing poverty—or does it lead to a decrease in their overall bargaining power within households, since they are no longer needed as much in production?

Wage effects may also lead to significant migration effects as workers move from rural to urban areas. These moves are not necessarily driven only by wages, and they may trigger poverty impacts that are distinct from wage effects. For instance, people moving from rural areas to urban ones may exchange informal risk sharing and safety nets (based on family and community) for formal programs that rely on public commitments of resources. The poverty effects are ambiguous but may be important.

2.3.6 Impacts on agricultural workers: non-wage effects

Although wages are the main channel of poverty impacts on agricultural workers, there are other possible pathways to poverty impact. As noted, migration decisions may be shaped by the demand for labor—and in turn be affected by new agricultural technologies. Other technology-related pathways to poverty impacts for agricultural workers may include issues such as exposures to chemicals and health risks. Workers exposed to toxic production conditions may end up impoverished owing to health expenses and time lost to work (and the time of family members lost to caring labor). In the same vein, agricultural technologies that reduce exposure to toxins or to accidents may have significant indirect impacts on poverty.

2.3.7 Impacts on non-agricultural workers

The poverty impacts of agricultural research on nonagricultural workers are almost by definition indirect in nature. The main effect on nonagricultural workers, leaving aside price effects that they may face as consumers, is likely to come through labor markets and to take the form of a wage effect. One main force affecting the wages of nonagricultural workers is the potential for an agricultural innovation to cause a change in the supply of workers in the nonagricultural sector. For instance, a new technology that displaces labor from agriculture is likely to increase the supply of labor in nonagriculture, thereby depressing wages in this sector. A second effect may come through changes in the demand for labor, perhaps driven by changes in rural demand for nonagricultural goods. For instance, a newly prosperous rural population may demand more goods and services, thereby increasing wages in the nonfarm economy (as in the classic Johnston and Mellor models). These are clearly second-order effects in terms of poverty impact, but in some settings these may be quantitatively important.

2.3.8 Impacts on landowners

Many productivity-enhancing technologies will increase the returns to land and will thus benefit landowners, who may or may not be farmers themselves. The class of landowners also typically includes people who own widely varying quantities of land. The distribution of landholdings may be highly consequential for the poverty impacts of agricultural research. Moreover, gains in the returns to land are not likely to be uniform across all types of land. In fact, in some circumstances, positive income effects for one group of landowners may be accompanied by negative income effects for other landowners—e.g., those who did not benefit from the new technology. For instance, a technology that favors rice production in one ecology (e.g., irrigated lowlands) might increase land rents in those areas while reducing them in other rice-growing areas. This differential impact of technology will affect producers, as mentioned, but it will also affect the value of land and the owners of land, who may or may not be farmers.

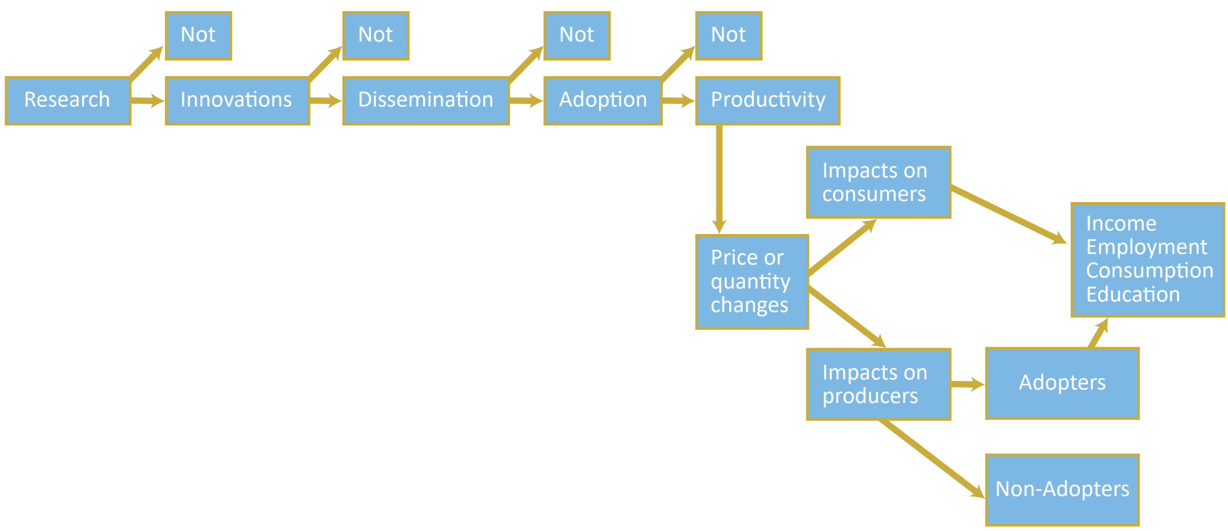
New technologies may also alter patterns of landownership, leading to the consolidation or splitting of landholdings; both patterns have been observed in different times and places. For instance, a new technology that creates incentives for land grabbing—by either domestic or foreign interests—may end up displacing those with weak political and legal claims to land, typically the poor and powerless. There have been claims in the past about technological innovation leading to this kind of expropriation of the poor (e.g., during the Green Revolution), and there are current concerns over similar patterns emerging in parts of Africa and Latin America. But equally, other changes in technology have been scale-neutral or undermined the advantages of large farmers.

2.4 MAKING SENSE OF PATHWAYS TO POVERTY IMPACT

The preceding list of pathways is not intended to be exclusive—but simply to point out that there are many different pathways from research to poverty impact. Not all of these pathways involve productivity gains; there are many other potential paths to poverty impacts. The analysis looked at the effects on different population groups, but it is worth noting again that each of these categories is heterogeneous, with poor and rich people in each group. The within-group effects may be at least as important as between-group effects (for example, with technologies that benefit large landowners at the expense of smallholders).

It is also important to note again that the poor may be spread across many categories and that many individuals belong to multiple categories, making it difficult to estimate poverty effects very accurately from models that deal only with aggregate categories. The complexity of the pathways to poverty raises enormous methodological challenges for impact assessment. Identifying a particular effect is difficult, and finding the entire net effect requires thinking through a wide range of possible effects, across multiple pathways.

Figure 1. The long causal chain from research to impact



Source: R.W. Herdt, personal communication.

3 METHODOLOGIES FOR ASSESSING POVERTY IMPACTS

In this section, we discuss different methodological approaches to assessing poverty impacts of agricultural research. We begin by discussing four empirical problems of impact assessment. We then summarize an extensive review of related recent literature and discuss the advantages and disadvantages of different methods that have been used previously in the impact assessment literature. To date, relatively few studies have actually attempted to estimate the poverty impacts of agricultural research. Although there have been numerous studies of technology adoption, and many more on the yield or productivity effects of new technologies, poverty impacts have been harder to quantify.

Understanding the impact of agricultural research on poverty is inherently challenging for a number of reasons, not the least of which is the long and complex pathway between research and its potential effect on alleviating poverty, as already discussed. In addition, there are a number of methodological problems related to (1) obtaining data that are representative of the population of concern, (2) accounting for the reality that people choose to adopt or not to adopt innovations and hence need for the statistical analysis to recognize such self-selection, (3) establishing a robust reflection of the counterfactual—what reality would have been for the population if the innovation had not been available, and (4) attributing an innovation to a particular source of research—in the case of primary interest to SPIA, a CGIAR-funded research activity.

3.1 EMPIRICAL ISSUES IN IMPACT ASSESSMENT

3.1.1 Representativeness

Representativeness of data requires careful understanding of the population of interest—in poverty studies, that is "the poor." Hence, data used as the basis for poverty impact assessment should represent the poor in a country, region or village. But most impact studies focus on the land area where an innovation is used, not on poor people. The areas chosen for study are often in regions selected because the innovation of interest seems to have been adopted there first. It is difficult to determine the likely impact on larger regions or the larger population of poor farmers or consumers from such studies. Most empirical studies done on the Green Revolution were based on data from quite unrepresentative samples—not only were they small samples, but they were biased by virtue of being chosen for the convenience of the researcher rather than to represent a specific population.

Of course, samples representative of the households in regions larger than villages or districts may get very large and entail costs beyond the resources available to most impact studies. This makes it doubly important to carefully define the population from which samples are selected and attempt to place them in the context of general interest—something like "the poor producers of the commodity in the country."

Stratified sampling is often used in an effort to address this challenge, but stratification can be overused or improperly used. For example, stratifying by region, district, sub district, and village based on where an innovation is reported to have been adopted leads to a highly selected set of villages. Choosing a random sample of households within such villages may give valuable data for areas of intense adoption but will give limited insights into why poor farmers in some regions adopt while in others they do not. Likewise, parameter estimates derived from such samples cannot be representative of what might happen if the innovation eventually enjoys widespread adoption. Careful thought is required to design sampling that represents the population(s) of interest, and equally careful consideration is needed when drawing conclusions or using the results for generalization.

In recent years the advent of nationally representative surveys of farmers or rural families, such as the World Bank's Living Standards Measurement Study (LSMS), has provided an opportunity for impact studies that potentially overcome some of those limitations. Working with the LSMS has its own challenges, but it does address the population sampling issue.

3.1.2 Self-selection or endogeneity

The fundamental statistical problem in understanding the impact of adopting research results is one that has been widely recognized in economics, and it pertains to a wide range of policy analysis and program evaluation. The basic problem is that people are not passive experimental subjects when it comes to real-world programmatic interventions: they are instead very likely to make choices that respond to the programs that are being implemented. This means that a researcher who comes along after some program has been enacted and tries to infer the impact of the program by comparing participants and nonparticipants will have difficulty extracting the "true" program effect. This program effect will be contaminated by a variety of other effects: participants may select in (or out) of program participation, based on whether or not they find it advantageous; they may also differ in other respects that are pertinent for program evaluation.

The statistical challenge is thus different from one that an agricultural scientist might face in looking at the effects of a particular agronomic treatment—such as in a fertilizer trial or a varietal trial. In that context, the researcher can assume that experimental controls have eliminated essentially all meaningful variation between treatments, so that the remaining differences are causally related to the treatment.

In impact evaluation, however, this is typically not the case. With the exception of randomized controlled trials (RCTs), discussed below, many or most impact evaluations rely on comparisons that are made *ex post* and in which a "treatment" group is in some fashion being compared with a "control" group. But the researcher is typically not able to assume that the treatment and control group differ only in their exposure to the program. Only in the case of experimental and quasi-experimental studies can this assumption be made—and, the treatment and control are often imperfectly distinguished.

This impact evaluation problem is at its heart an issue of the potential endogeneity of the treatment variable. Where data are observational, exposure to technology is not randomly assigned. A farmer's decision to adopt a new technology is likely to be influenced by observable and unobservable variables. The unobservables might include farmers' motivation and managerial skills, quality of land, and other factors. When these unobservables are correlated with the outcome of interest, it can bias the estimation of treatment effects. If the treatment is a new crop variety, for instance, it is highly unlikely that farmers take it up randomly. A comparison of yields between those who are using the new variety and those who are not using the variety is likely to reflect the underlying differences between the two groups—not just the effect of the improved variety.

For poverty impacts, the problem is particularly acute. A concrete example is useful: Suppose we find that in a particular location those farmers who have adopted improved varieties have higher incomes and experience lower rates of poverty than those farmers who do not use the improved varieties. The difference might be due to the varietal choice. But, the difference might also be due to other factors that are correlated with varietal choice. Suppose, for instance, that rates of adoption are higher among farmers with good land and access to irrigation—and perhaps also those with good access to markets. These farmers might have attained higher yields even with the pre-existing varieties. This means that we cannot infer that the yield differences can be attributed to the variety. An agricultural scientist might imagine that we could properly deal with this problem simply by controlling for irrigation, market access, and all the other production characteristics that vary across fields. And indeed, if all of these characteristics were observable and were quantifiable, it might be possible to extract the statistical effect of the variety. But if there are also unobservable differences between adopting and non-adopting farmers, or between their farms, then the statistical problem remains.

3.1.3 Counterfactual

The importance of establishing a robust, credible counterfactual, rather than simply describing events after the introduction of an innovation, has become recognized as a critical problem in impact analysis over the past two decades and is related to but reaches beyond appropriate statistical treatment of data. A valid impact assessment shows what the effect of an innovation has been on the variables of interest—that is, the difference between what the situation actually is and what it would have been in the absence of the innovation. Of course, the latter cannot be observed. In most cases, one can observe "adopters" and "non-adopters" of the innovation, but the two categories may differ in many ways other than their adoption of the innovation, so looking at non-adopters is generally not a robust basis for a counterfactual. In some cases it may be possible to have information on both adopters and non-adopters before and after the introduction of the innovation, providing a more robust basis for the counterfactual. To add to the complexity, policies, prices, and institutions may change after the introduction of an innovation, for reasons unrelated to the innovation, and affect the variables of interest. Statistical methods for addressing these challenges are discussed below, but careful thought about what likely would have happened in the absence of the innovation is essential for good impact analysis and helps in determining the statistical approach to take.

3.1.4 Attribution

Attribution, stating that an observed change was caused by a particular research activity, is a critical element in impact analysis but one that is fraught with pitfalls. The case of a genetic innovation, such as a crop variety, would seem to be simplest because one can identify the breeding organization that submits the variety for release. However, every variety is built on its parents and earlier ancestors, which generally come from diverse sources and perhaps can be traced back to landraces. Hence, while one may attribute some share of the contribution to the breeder of the variety, some share might also be attributed to earlier sources. In addition, the multiplication and dissemination of a variety require effort, expense, and in many cases creative thought, so some share may be attributed to dissemination. Innovations in management of soil and water may be more complex, and institutional or policy innovations even more so, as generally the intellectual effort involved in their creation is considerable and builds on earlier analyses. One has only to look at the references cited in most research papers to understand that tracing ideas to their origin is probably not only difficult but impossible. This all suggests that attribution cannot be accomplished with a quantitative tool but rather requires a careful narrative discussion of the pertinent details. CGIAR research seldom is responsible for more than a small fraction of observed productivity gains.

3.2 STATISTICAL APPROACHES

A variety of statistical approaches have been developed that attempt to deal with impact evaluation problems in different ways. This section reviews some of the recent literature on the links between agricultural research and poverty reduction. The full literature review appears in Appendix 1.

The review conducted for this paper was not intended to be comprehensive or to represent a systematic review of the literature. Instead, it was an overview of those papers most closely related to the subject of this paper, with the aim of providing a representative overview of the typical studies and their findings. Our review was limited to papers that have been published (in peer-reviewed journals or as working papers) during the past 10 years. We have further included the studies on this topic commissioned by SPIA.

We would be remiss if we did not acknowledge an important older literature—largely qualitative or descriptive rather than rigorously quantitative—on the social impacts of agricultural technology. Much of this literature has focused on the Green Revolution and its impacts, and a substantial amount of it was sharply critical of the impacts of agricultural technologies on poor people and their communities. A recent literature has also brought a critical lens to innovations emerging from the private sector, such as the introduction of Bt cotton in India and the diffusion of genetically modified organisms (GMOs). We touch on this literature below in our discussion of qualitative and historical methods of assessing poverty impacts. But, it is important at the outset to recognize an important lesson that emerges from the critical literature: not all agricultural technologies generate beneficial impacts on the poor. There is real potential for new technologies to harm the poor, and in many cases the impacts of new technologies will display strong heterogeneity across populations. It is thus important to opt for research designs that do not implicitly assume that technologies are always beneficial to the poor, and to be aware of the potential for highly productive technologies to prove harmful to the poor.

The main focus of our literature review, however, is on recent empirical studies. We consider 58 papers from the past decade. These can be subdivided into three main groups. The first and largest group (36 papers) consists of micro studies examining the link between modern agricultural technologies (which we view as outputs of agricultural research), their adoption, and their direct farm-level impacts.[1] These papers typically focus on one specific technology (e.g., an improved crop variety) in one country and use observational or experimental data for econometric analysis or for simulations within agricultural household models. The second set of papers that we consider is a small set (5 papers) of micro- and meso-studies examining the same relationship between modern agricultural technologies, their adoption, and a set of slightly broader outcomes—such as indirect impacts on the overall population. Finally, the third group (17 papers) comprises macroeconomic papers that examine the impact of general agricultural productivity growth on poverty reduction. Most of these rely on model-based simulations, such as computable general equilibrium (CGE) models at the country level or similar models of international markets (e.g., those based on the multi-country Global Trade Analysis Project [GTAP] framework). A few use econometric analysis.

3.2.1 Micro studies linking agricultural technology to direct farm-level outcomes

Micro studies investigating the impact of agricultural technology on poverty rely mainly on obser-

[1] Note that the sample is skewed in the sense that, with a few exceptions, we did not look at papers that solely focused on adoption. Instead, we defined our sample to include studies that assessed at least one outcome beyond adoption, such as yields.

vational data (usually cross-sectional, sometimes longitudinal) from household surveys. A large number of these studies have been carried out by researchers based in the CGIAR system, often relying on surveys designed and conducted specifically for the purpose of documenting adoption and impact on farm households. In recent years researchers have also increasingly made use of data obtained in experimental settings, i.e., randomized controlled trials. Although the number of published studies using such experimental data on agricultural research impacts is still limited, the number of currently running trials suggests this part of the literature is likely to grow in the near future.

As discussed below, each method has advantages and disadvantages for addressing certain questions, and requires making assumptions that are reasonable and defensible for the particular method, context and data chosen. It is important to be very explicit about the assumptions made and why they are reasonable, and to provide supporting evidence when possible.

3.2.1.1 Studies using observational data

The studies that we considered focus largely on countries in sub-Saharan Africa: Ethiopia (four), Kenya and Uganda (three each), Malawi and Zambia (two each), and Mali, Mozambique, Nigeria, and Tanzania (one each). Only a few studies have looked at Asian or Latin American countries. As mentioned, these papers usually investigate the impact of a single technology. In about two-thirds of the papers, this technology is an improved/modern crop variety. The largest number of studies addresses maize varieties, but other studies look at improved varieties of wheat, rice, sorghum and millet, groundnut, beans, pigeonpea, and Bt cotton. A few studies also evaluate the impact of extension, irrigation, or soil conservation measures. Most papers look at more than one outcome: beyond technology adoption, the most frequent outcomes assessed in these papers are yields (about half of papers), some measure of household income or expenditure (two-thirds), food security (one-fourth), and saving and assets indicators (one-fourth), as well as poverty (head count, gap, or severity) and inequality measures (one-fourth). Two-thirds of studies further investigate the heterogeneity of impacts: the dimensions most frequently used for this are farm size, income/wealth, and education. Single studies also disaggregate based on actual or potential yield quintiles, gender of household head, adoption propensity, or fertilizer subsidy receipt.

The main methodological challenge of these studies is dealing with the potential endogeneity of the treatment variable. The approaches to deal with this problem depend on the type of data available (e.g., cross-sectional versus longitudinal). In papers entirely relying on cross-sectional data, the two main methods employed are instrumental variables–based approaches (either two-stage least squares [2SLS] or Heckman selection-type models) and propensity score matching (PSM). A few papers use both of these approaches (either in combination or one as a robustness check for the other).

3.2.1.1.1 Instrumental variables

Instrumental variables (IV) methods aim to resolve the problem of the endogeneity of the treatment variable by means of an instrument that is correlated with the treatment variable (relevance condition), but uncorrelated with the error term in the outcome equation (i.e., the unobserved factors that affect the outcome, validity condition). The variables used in the literature to instrument for adoption in the outcome equation are of three main types:

1. the financial and/or transaction cost the household must incur to access seeds (e.g., the distance to nearest seed source/seller, droughts and floods in the past 10 years, the seed price or the seed-to-grain price ratio, the number of years a household has been receiving a maize subsidy, or the existence of credit service);
2. some measure of remoteness or the existing market infrastructure (e.g., the existence of marketing service for agricultural crops in the village, distance to the main market, the quality

of roads to the main market);
3. the information on the new technology available to the farmer (e.g., the distance/access to the agricultural extension office of a governmental or nongovernmental organization or a farmer cooperative, the cumulative adoption rate in the village or within the social network, participation in some form of training or participatory variety selection).

In principle, the same approaches could be used to identify poverty impacts instead of yield effects or other treatment effects related to improved technologies. But there are strong reasons to be wary of the IV approaches used here.

Instruments of the kind that have been used in the literature are likely to be correlated with adoption, and hence are relevant and potentially even relatively strong. At the same time, almost all of these are likely to be correlated with unobserved factors in the outcome equation, meaning that they are correlated with the outcome variable through channels other than the adoption decision. Hence, their validity may be questionable. For example, unmeasured farmer motivation or ability could affect the farmer's adoption decision and simultaneously result in higher-than-average yields, independent of the crop type grown. This will depend on the specific context, so no general judgment can be made. But it is clear that a thorough justification of the validity of the instrument, based on economic theory or some supplementary data, is crucial. It also illustrates that instruments that may be very useful in certain contexts may not be appropriate in other contexts. This may appear self-evident; nevertheless, in the literature reviewed here, instruments are justified far too frequently based on the fact that (1) they have been used similarly in other published papers, (2) they pass tests of over-identification or exogeneity, or (3) the coefficient on the instrument in the outcome equation is insignificant.[2] None of these justifications should be viewed as adequate. Instead, it should be obligatory for researchers to explain why the instruments satisfy the validity criteria.

3.2.1.1.2 Selection models

Additional concerns apply to the class of 2SLS-based estimators and Heckman selection models. These models generally estimate a single outcome equation for adopters and non-adopters. Implicitly, they assume that adoption only results in an intercept shift, so that the coefficients on the other covariates are the same for adopters and non-adopters. To relax this assumption, a few papers use endogenous (or occasionally exogenous) switching regressions—i.e., they estimate separate outcome regressions for adopters and non-adopters. However, the switching regressions rely on a more restrictive exclusion restriction along with the less-than-innocuous assumption that the error terms of the selection equation and the two different outcome equations follow a joint trivariate normal distribution.

3.2.1.1.3 Propensity score matching

Propensity score matching (PSM) offers a different approach to dealing with the problem of identifying treatment effects given the endogeneity of treatment—i.e., the fact that adopters and non-adopters differ in many ways other than their adoption decision. PSM approaches assume that the heterogeneity correlated with the outcome of interest is fully observable. It is thus possible to match households based on observable characteristics, so that only the treatment will vary. Ideally, the observables should influence both the treatment assignment and the outcomes but should not be affected by the treatment—i.e., the causation flows in only one direction. To better deal with the multidimensionality of these characteristics, matching is performed using a propensity score. This propensity score is estimated as probit model (rarely as a logit model) where adoption is regressed on all potential household or village characteristics that could influence adoption. Based on this propensity score, adopters and non-adopters are matched using nearest-neighbor or kernel-based methods. (Most papers present estimates for both approaches.)

[2] As pointed out by Murray (2006) and Sovey and Green (2011), this is not a valid way of examining the validity of an instrument since this regression includes the endogenous variable, and hence the coefficient on the instruments will be biased.

Thus, matching is essentially used to trim the sample of households that are compared.

As previously stated, the core assumption on which propensity score matching relies is that of conditional independence or selection on observables, i.e., that once all observables have been controlled for, the adoption decision is not correlated with the error term. This is also the assumption this approach is most criticized for: assuming that adoption is a rational economic decision of farm households, adopters and non-adopters that share the same observable characteristics have to systematically differ in their unobservables, thus violating the conditional independence assumption (de Janvry et al., 2011). However, it is to some extent possible to evaluate the impact such unobserved heterogeneity will have on the PSM-based estimates by calculating Rosenbaum bounds, an exercise performed by an increasing number of papers. Obviously, the assumption of selection on observables is easier to defend if the technology diffusion mechanisms/patterns followed a roll-out/placement based on observable characteristics (e.g., members of farmer associations were targeted in certain regions but not others and targeting criteria was explicitly defined) and this can be documented with available data.

Like IV-based methods, simple PSM assumes homogeneity in the impact of adoption; in other words, it is assumed that the effect of adopting a new technology is the same on all farm households. To relax this assumption, some studies examine whether there is heterogeneity depending on the propensity score, analyzing whether farmers who are more likely to adopt (according to the propensity score) benefit more or less than farmers less likely to adopt.

3.2.1.1.4 Panel data

Eight papers in our review make use of panel data that span a mean time period of 7.6 years and include an average of three waves. These papers mainly use one of two econometric approaches: difference-in-differences or household fixed effects (FE). Some of the papers also use an instrumental variables approach. FE estimators eliminate any potential bias stemming from time-invariant unobservables, yet they require that the treatment/adoption variable varies sufficiently over time. In cases where the data do not exhibit sufficient variation in the treatment variable or where the dependent variable is non-linear (e.g., with some poverty measures), papers resort to correlated random effects (CRE) estimators. However, these come at the expense of assumptions that are far more restrictive. A key assumption of panel data is that of a common trend, and again data can and should be used to show common pre-trends in the outcomes of interest.

3.2.1.2 Studies using experimental data

As mentioned, the number of impact studies of agricultural research using experimental data still remains limited. Like the observational studies reviewed in the previous section, the eight experimental papers included in this review mainly focus on sub-Saharan African countries (the Democratic Republic of Congo, Kenya, Mozambique, Rwanda, Sierra Leone, Uganda, and a multi-country panel of 8 sub-Saharan Africa countries). In addition, two other studies use data from India. In terms of the experimental interventions assessed, two papers center on improved rice varieties (flood-resistant Swarna-Sub1 and NERICA), two look at improved maize and fertilizer, two sister papers look at "innovation platforms," and one paper evaluates mobile phone–based extension services. Half of the papers randomize at the individual level only; the other half randomize at both the village and the individual level. All eight of these experimental trials carried out a pre-intervention baseline, followed by at least two follow-up surveys, usually one and two years after the intervention.

The outcomes investigated are similar to the observational studies. Three of the papers look at technology adoption as their outcome measure. Four look at crop yields, and two each consider various measures of consumption, food security, savings, and assets. One paper reports measures of poverty head counts. Additional outcomes evaluated are farmer decision-making (two papers), farmer knowledge, rate of return to fertilizer, adoption of credit, and housing improvements (one paper each). As with the non-experimental papers, heterogeneity of impacts is disaggregated by farm size, yield distribution, and level of education, as well as by risk aversion and number of households treated in the social network. Finally, none of the papers reports results on differential impacts across the wealth/income distribution.

Experimental studies aim to resolve the problem of causal inference by randomizing assignment to treatment. By construction, the two groups will thus be drawn from the same population such that treatment status is orthogonal to baseline observable and unobservable characteristics. Computing the difference-in-means (or if a baseline is available, the difference-in-differences) between treatment and control groups then yields the intention-to-treat (ITT) estimate, which under perfect compliance is equivalent to the average treatment effect on the treated (ATT). Perfect compliance implies that everybody who is assigned to the treatment group actually adopts the technology, but nobody in the control group does so. Research design should start from key identifying assumptions and provide empirical support for their plausibility.

As a practical matter, most of the RCTs in this review achieved only partial compliance in adoption: Swarna-Sub1 adoption is 76 percent among treated households against 10.1 percent among control households in the same village (Emerick et al., 2014). The partial input subsidy voucher program in Mozambique achieved a 43 percent adoption rate among the treated against 12 percent among non-treated (Carter et al., 2013). Imperfect compliance appears to be more pronounced if seeds or fertilizer are sold at market price or if only a partial subsidy is granted. But this pricing approach is frequently practiced, in part to avoid so-called Santa Claus effects. Similarly, collaboration with local government agencies results in leakages of subsidies or program management that undermines random assignment (Emerick et al., 2014; Duflo et al., 2009; Carter et al., 2013).

Partial compliance is particularly problematic when treatment effects are heterogeneous, as might typically be expected with the adoption of a new agricultural technology. In this case, the ATT estimate (i.e., the impact of the treatment on those in the treatment group who actually adopted) is likely to suffer from selection bias. Because of this problem, studies with substantially imperfect compliance normally only report intention-to-treat (ITT) effects; i.e., they compare all individuals originally assigned to the treatment group with those originally assigned to the control group, regardless of the actual take-up of the technology. From a policy perspective, this ITT estimate actually may be the more relevant measure, since it reflects the potential impact a scaled-up intervention could achieve on the whole population. But the ITT estimate is, in some sense, a watered-down version of the ATT.

Another concern arises from the fact that in all studies, the group of farmers that participates in the randomized trial is already a selected group of farmers that has to fit certain pre-defined criteria (interest in participating in the trial, specified farm size, etc.). While this may help to increase compliance, it also entails the risk that the trial may measure the impact of the technology on a specific subset of farmers, but not the whole distribution of farmers. Thus, a trade-off may exist between the degree of compliance achieved and the heterogeneity of farmers.

3.2.2 Studies evaluating direct and indirect effects of adoption of a new agricultural technology

The previous section summarizes some studies that link agricultural technologies to farm-level impacts. A small literature has sought to go one step farther, linking technology adoption to various measures of broader impacts. Our review of the SPIA-related literature found a small set of studies that used micro-econometric results to compute effects beyond the farm level. In contrast to the studies discussed in the previous section, these all attempt to locate the economic impacts of new technologies on groups other than the farm households that use them. These studies are thus able to shed light on a different set of the pathways to poverty impact. It is worth noting that finding poverty impacts on producers from a price-quantity effect is often a tall order even if there are large productivity improvements of one variety/practice, just because that one variety is often a tiny fraction of the farmer's diversified income portfolio. Thus, there is an inherent lack of power that will affect almost any micro study looking at poverty impacts of single innovations on producers.

Three papers in this group provided estimates of the economic surplus generated by an improved variety—one each for maize (Zeng et al., 2013), beans (Larochelle et al., 2013), and rice (Raitzer et al., 2013). Although these papers used slightly different approaches, the basic idea was to estimate supply and demand curves and then to attempt to quantify the shift in the supply curve brought about by agricultural innovation.

A different approach was taken by Subramanian and Qaim (2010), who use a micro-social accounting matrix (SAM) from a census survey in an Indian village to simulate the impacts of extending the area under cultivation of Bt cotton on different types of households. Finally, Minten and Barrett (2008) use cross-sectional commune-level data to econometrically investigate the relationship between higher rice yields (assumed to reflect improved technology) and a set of outcome variables that are directly linked to poverty impacts: real wages for unskilled labor, staple food prices, and welfare indicators.

In all cases, the challenge of these approaches is (1) to develop convincing micro-econometric evidence on the key parameters of the model and (2) to write down models that are appropriate and useful depictions of reality. The three economic surplus studies rely on estimates of the yield impacts of new varieties in order to compute the downward or pivotal shift in the aggregate supply curve that may be attributed to the improved variety. The studies reviewed here estimate the impact of the improved variety on yields and costs based on production function regressions (analogous to those used in observational micro studies). These are estimated in some cases using cross-sectional data (and different IV-based estimators), and in other cases combining household and municipal/province-level panel data (and FE estimation).

The impact of the improved technology on the economic surplus further depends on the assumption regarding tradability of agricultural goods—i.e., whether a closed or a small open economy (SOE) is considered (Maredia et al., 2000). In the case of an SOE, prices will remain unaffected and the whole additional surplus generated by the improved variety will be allocated to producers. In contrast, in a closed economy, the shift of the supply curve will result in lower prices and the additional surplus will be split between producers and consumers. In this latter case, further assumptions on the price elasticity of demand and supply are necessary to determine the relative distribution of surplus between producers and consumers. Assumptions on tradability also vary between the three papers: one considers a small open economy, one a closed economy, and the third compares both cases. Both studies that examine a closed economy case employ constant elasticity functions for both demand and supply; one of them further compares this with a case with a positive supply shutdown price. To obtain the final impact on poverty, this study then allocates the changes in surplus to households depending on their net sales position. The study by Raitzer et al. (2013) computes health (due to nutrition impacts) and environmental impacts (mainly due to prevented deforestation) as additional outcomes.

Although these models have the advantage that they go beyond the direct farm-level impact studies to look for broader impacts, they face the same the methodological issues as the narrower studies because the empirical estimation still underpins the models that are used. The issues of causal identification remain, and they are compounded by questions about the correct modeling of demand and prices. The poverty impacts calculated with models of this kind are ultimately somewhat mechanical. The economic surplus studies are based on the price-quantity pathways discussed above, and they cannot capture other pathways very effectively.

3.2.3 Macro studies on the impact of agricultural productivity growth on poverty reduction

A number of papers use macro approaches, rather than micro approaches, to examine the relationships between agricultural productivity gains and poverty reduction. The largest part of this literature uses econometric methods to estimate the elasticity of poverty with respect to agricultural research and/or agricultural productivity increases. A second group of papers relies on model-based approaches (chiefly computable general equilibrium models but also some multi-sector growth models). These mostly attempt to simulate the consequences of a productivity increase in agriculture on aggregate income growth as well as poverty. In both types of studies, there is normally a comparison to productivity change in other sectors. For instance, several studies seek to compare agricultural productivity growth effects with non-agricultural growth effects. Outcome variables of interest usually include both direct and indirect channels (e.g., wages and food prices) through which effects may operate.

3.2.3.1 Regression-based macroeconomic studies

The regression-based studies that we reviewed mainly use multi-country macro-level panel data in simultaneous equation models (e.g., estimated with three-stage least squares [3SLS] on the pooled dataset) in order to obtain the elasticity of poverty to agricultural research and/or productivity growth in agriculture. A few papers instead use simple ordinary least squares (OLS) or 2SLS. The most frequently used cross-country data are poverty data assembled by the World Bank. Other sources include poverty data from the International Labour Organization; IFPRI data on agricultural research; and sectoral data from UN national accounts. In countries where large longitudinal household survey datasets are available (China, Ethiopia, and India), authors also use single-country panel data. Across studies, it is difficult to compare results because of differences in the growth concepts employed (e.g., agricultural labor productivity growth, sectoral value added) and in the range of outcome measures employed (e.g., household expenditure, poverty headcounts, or poverty gaps) (de Janvry and Sadoulet, 2009).

Econometrically speaking, these studies face a number of problems. The chief issue is that the right-hand variables in these regressions are not in any sense randomly allocated. The relationships in the data that result may reflect reverse correlation, spurious correlation, and simple specification bias, as well as causal relationships. For instance, a positive correlation between agricultural productivity growth and agricultural research may reflect an underlying causal relationship. But it may also be the case that research expenditures are likely to rise in a growing economy. Or it may reflect the fact that good government policies typically lead both to increases in research spending and to declining poverty, without any necessary causal relationship between the two. Econometric panel methods can provide some ability to address these problems (e.g., through the use of time lags and the imposition of structural restrictions on the data), but few of the papers in our data made use of these approaches.

Another issue that arises, particularly in the within-country studies, is the need to control for spatial patterns of economic activity and for other kinds of spillovers. The basic problem of statistical inference here is that the different observations (e.g., on neighboring districts) are not fully independent. Because of cross-location spillovers, productivity growth in one district is likely to affect what happens in nearby districts or those that are otherwise linked through trade. Again, there are tools from spatial econometrics that can address these statistical problems with varying degrees of success, but these tools have not been widely used in the literature.

Perhaps a more serious concern with this literature is that it has often ignored the differences between poverty-growth elasticities or poverty-productivity elasticities and measures of returns to investment. As argued in Dercon and Gollin (2014), the poverty-reducing effects of agricultural growth compared with other sectors do not necessarily imply that investments in agriculture will have greater poverty impacts than investments in other sectors. It will depend on how readily investments translate into growth in different sectors. If it is particularly difficult to generate growth in agriculture, then even the high poverty-growth elasticities will not necessarily imply that agricultural investments have a high poverty impact.

Another concern is that, from a methodological perspective, these models struggle to explain the exact mechanism or pathways through which agricultural productivity growth leads to poverty reduction. Because they typically use aggregate data and focus on fairly broadly defined outcome measures, it is not always clear why or how poverty reduction is being achieved. One study (Datt and Ravallion, 1998a) offers some insight, concluding that general equilibrium effects through higher wages and lower prices appear to be responsible for most of the beneficial effects of agricultural productivity growth in India. Even small changes in food prices appear to have considerable effects on absolute poverty. This finding is certainly consistent with theory, but the results may be quite specific to India, where many of the poorest people are landless workers, in either agriculture or non-agriculture.

3.2.3.2 Model-based macroeconomic studies

Several papers approach the challenge of poverty impact assessment from a methodological perspective that is based on large single-country or multi-country macro models. These models are typically partly or fully general equilibrium in flavor. Many are CGE models built around a core input-output structure, in the form of a SAM. A virtue of these models is that they have a mechanism for capturing the economy-wide effects of an agricultural productivity change or technological innovation, including effects that occur outside the agricultural sector. Many of the models also include a large number of household types, allowing for the kind of disaggregated analysis that is valuable for poverty impact study.

In general, the multi-country models are organized around world markets and are calibrated to match observed patterns of production and trade for a subset of commodities. These models are typically constructed in such a way as to allow for modeling of changes in policy or productivity in some subset of countries. The basic model framework can then be modified to address a wide range of questions. For instance, the GTAP model—one of the most widely used multi-country models—was originally developed to model agricultural trade liberalization, but has subsequently been modified to allow for analysis of a range of environmental and labor issues, as well as poverty analysis.

The single-country CGE models offer more detail on the input-output structure of a single economy. The models can accommodate a large number of production sectors and household types, allowing for elaborately detailed structures. Various additional dimensions can be grafted onto these models, so long as they are incorporated in an internally consistent way. By construction, the models replicate the social accounting matrix on which they are calibrated.

Poverty impacts in these models are usually estimated by looking at the welfare of a set of households who are defined initially as poor. In some (but not all) frameworks, the number of people who are poor will change; in others, the size of this group is taken as a given and only their welfare changes. The methodological challenge is that results of model experiments are difficult to validate.

These models offer powerful tools. They are internally consistent, and they can be used to conduct a variety of experiments that would be difficult or impossible (or unethical) to carry out in the real world. Within these model economies, it is normally straightforward to ask how a particular policy change or productivity improvement will alter the well-being of poor households. The models, depending on their level of detail, can address some very specific questions. For instance, researchers could ask whether an increase in cassava productivity will have larger poverty impacts than an increase in maize productivity. As such, these models can provide a valuable tool for researchers interested in counterfactual scenarios of various kinds. Causal identification is clear within these models; a particular change can be said to generate a causal impact on other variables.

The methodological challenge of these models is that the calibration often requires strong structural assumptions about functional forms and model specification. For instance, the substitution relationships in the models (e.g., consumers' willingness to substitute across goods, or producers' willingness to substitute across inputs and outputs, or the substitutability of imports for domestic goods) are often important for the results. Often, however, the data provide little information on the underlying substitution elasticities. Other aspects of the models may be similarly important for results and may require parameter values that are equally difficult to identify in the data. Moreover, the sheer complexity of the models makes it difficult to know how sensitive they are to particular assumptions. The more sectors and the more structure the researcher adds to the model, the less straightforward it is either to validate the model or to understand its sensitivities.

Thus, although the models are strong on internal consistency, they can be weak on external validity: it is difficult to know how much they behave like the economies that they represent, other than replicating exactly some set of baseline observations. Seldom do researchers validate the models by using them for retroactive forecasting ("backcasting") or for examining real-world policy "experiments" for which the outcomes are known.

3.3 QUALITATIVE RESEARCH

Although our literature review has prioritized empirical studies with a quantitative dimension, an extensive qualitative literature has also examined the impacts of agricultural research on poverty. In particular, a number of studies, particularly from earlier years, focused on the impact of the Green Revolution on poverty and inequality in South Asia. Since we chose to focus on more recent (and quantitative) papers, this part of the literature has not been included in the systematic review. Furthermore, a comprehensive review of this literature can already be found in Hazell (2010) and Kerr and Kolavalli (1999).

Qualitative approaches range from purely narrative accounts to those that gather extensive descriptive data, and even to qualitative-quantitative ("Q-squared") approaches. Because qualitative approaches can take a broad view of impact, they can provide rich and detailed accounts of how new agricultural technologies alter the conditions under which individuals live and work and how innovations change social structures and dynamics. Qualitative studies can offer a richness of detail that is not available in quantitative analyses.

Arguably, some of the best descriptions of the long-term impact of improved agricultural technologies come from long-term mixed-method studies such as those by Yujiro Hayami, Masao Kikuchi, and colleagues in the East Laguna village in the Philippines (Hayami and Kikuchi, 2000), or the long-term studies carried out by Nicholas Stern, Peter Lanjouw, and others in Palanpur, an Indian village in the state of Uttar Pradesh (Lanjouw and Stern, 1998). These studies follow particular village locations over time, during episodes of rapid technological change in agriculture, and they are able to capture the evolution of the respective village economies in great detail. Through repeated resurveys in combination with broader social science research, the studies provide deep and detailed understanding of the process of development as it has unfolded over five decades.

Other qualitative approaches may involve much shorter timescales. Methods such as interviews, anthropological observation, historical analysis, and other approaches can reveal striking truths about the process of agricultural technology adoption, diffusion and impact. In many cases, qualitative work has paved the way for subsequent quantitative work that has confirmed (or occasionally challenged) hypotheses that emerged from the qualitative research. Used in combination, qualitative and quantitative work can provide valuable insights into the processes of poverty reduction, as argued by White (2002), Kanbur and Shaffer (2007), and Shaffer (2013).

Some of the challenges with using qualitative work for poverty impact assessment are similar to those that emerge with quantitative studies. One difficulty is establishing clear and convincing counterfactuals. To understand the impact of agricultural technologies on the evolution and development of a community or group of individuals, we need some way to think about what would have happened in the absence of the technology. Such a counterfactual scenario may be as elusive for qualitative researchers as for quantitative researchers; comparisons across communities or over time are likely to be confused in precisely the same ways by selection bias, unobserved variation, and the potential for reverse causation. For the same reasons that we question quantitative studies that compare adopters with non-adopters, or communities with and without improved technologies, we should question the same comparisons when they involve qualitative methods.

Qualitative studies are also limited in many cases by the trade-offs between geographic scope and depth. Most studies that use anthropological approaches are limited to narrow geographic areas that may or may not be representative of larger settings. And while long-term historical analyses can provide richly textured descriptions of economic and social change, the complexity of historical processes is such that it can be very difficult to arrive at any clear causal understanding of the impacts of technological change. For instance, it can be challenging in a study such as the Laguna village surveys to identify the effects of new agricultural technologies separate from the effects of other changes taking place over time—in education, in urban markets and transportation, in policies and political structures, and in climate.

3.4 OTHER METHODS NOT APPEARING IN THE LITERATURE

Our literature review did not find any examples of two other methods that should in principle be valuable in contributing to the analysis of poverty impacts. One is the use of natural experiments and quasi-experimental designs; in principle, these seem like promising avenues to explore. The other is the use of local economy-wide impact evaluation (LEWIE) models and approaches.

3.4.1 Natural experiments and quasi-experiments

In the development economics literature, there has been growing use of "natural experiments" and quasi-experiments as tools for impact assessment. The idea of natural experiments is to find exogenous shocks or events that lead to plausibly exogenous variation in treatment across groups or locations (or, in some cases, exogenous variation in the intensity of treatment). For instance, in the case of CGIAR, the decision in the 1960s to do research on barley but not on oats might suggest a plausible research design; similar decisions led CGIAR to conduct research on lentils but not mung beans, and so on. In principle, it might be possible to ask whether these decisions led to different outcomes between areas suitable for different crops. Similarly, we could imagine that exogenous variation in soils or geography proved critical in determining the suitability of different locations for improved crop varieties or animal breeds. By comparing locations that differ only (or almost only) in these exogenous characteristics, it may be possible to treat the outcomes as the result of a "natural experiment." This approach would include a number of "regression discontinuity" methods that are similar in spirit to the natural experiments. Examples of these methods applied to agricultural technology impacts are recent work by Hornbeck and Keskin (2014) and Bustos et al. (2013).

We could not find examples of natural experiments or regression discontinuities used in impact assessments of CGIAR technologies. Since these designs may allow for the exploration of larger-scale effects than RCTs or other controlled experiments, they are potentially quite valuable for the kinds of impact assessment that are of interest to CGIAR. In addition, because many technologies may diffuse along fairly strong agroecological gradients, with well-defined domains, it may be possible to find good sources of exogenous variation based on differences in soils and climate. The lack of studies using these methods is striking, and the potential for exploration is high.

3.4.2 Local economy-wide impact evaluation (LEWIE)

LEWIE modeling is a form of project impact evaluation that focuses on the linkages within a local economy that transmit impacts from actors directly affected by a project/RCT ("treated") to other actors ("untreated") in the economy (Taylor and Filipski, 2014). Given the complexity of impacts, this method departs from partial approaches such as standard cost-benefit analysis.

LEWIE models were originally designed to evaluate the spillover effects of cash transfer programs, and they have their foundations in general equilibrium theory. They are somewhat related to macro models that use SAMs and CGE models, although they also incorporate ingredients from more micro-oriented agricultural household models and disaggregated micro economy wide models (Taylor et al., 2013).

The main element is a set of equations that describes the consumption and production behavior of households interacting in markets. Several groups of household exist: these differ in their activities, income mixes, consumption patterns, technologies used, and extent of integration in markets. Both as producers and as consumers, households respond to price changes. How prices are determined depends on the assumptions regarding the structure of local markets. Prices can be exogenous (if determined in outside or foreign markets), local (if a result of a village market-clearing condition), or household-specific (shadow prices in subsistence production households). Prices are further affected by transaction costs at the village and household level.

These models thus share many features with CGE models; however, they focus on a small regional economy with distinct but interacting agents (households). This implies that good knowledge of the local market structure is a key ingredient for successful modeling. One of the main challenges of LEWIE models is to strike a balance between complexity and feasibility: they should be complex enough to represent the most important interactions that transmit the impact of shocks, projects, and policies within local economies; at the same time they must be so simple that the necessary information can be realistically obtained either from existing data or through surveys.

LEWIE models are either constructed based on SAMs or calibrated directly from household survey data. The most challenging data requirement is usually data on the location of income and expenditure of households, which often are not part of standard household surveys.

In the context of evaluating the impact of agricultural technology on poverty, LEWIEs have not yet been extensively used. However, given the likely magnitude of indirect effects of agricultural technology on wages and prices and the numerous market imperfections confronting agricultural households and particularly the poor, LEWIE approaches can provide further interesting insights. This is illustrated in Filipski et al. (2013), which analyzes the potential impact an irrigation project in Tanzania could have not only on the actual project area but also on the region surrounding it.

3.5 SUMMARY

The literature review makes it clear that numerous quantitative and qualitative methods can be used appropriately for poverty impact assessment. All of the methods have corresponding pitfalls and limitations. It will be difficult to demonstrate poverty impacts with any single study using any particular method. But a set of related studies, focusing on a single technology or innovation, can collectively provide a convincing picture of poverty impacts. For instance, if a new technology is seen in an observational cross-section to produce higher profits for adopting farmers than for similar non-adopters (and especially if the result is robust to the choice of cross-sectional methods), this would be relevant information. If the same technology is shown, in the context of a country-level CGE model to generate beneficial poverty impacts, it would add to the strength of the evidence. And if, in addition, panel data show that poverty reduction is associated with the diffusion of the new technology, then at some point the evidence becomes compelling. Clearly, it is not possible to assemble this kind of evidence for each and every innovation emerging from CGIAR, but it would be valuable to develop this kind of evidence for a representative or indicative set of technologies.

As CGIAR moves forward, it might be wise to lay the groundwork for this kind of coordinated program of research, focused on a few technologies and innovations that are deemed likely to achieve poverty impacts. The effort to assemble this kind of evidence will require strong and coordinated planning at the level of individual research programs. It is not realistic to imagine that this kind of evidence can be produced *ex post* from the data that happen to have been collected; there will need to be a strong and forward-looking commitment to impact assessment, along with a sustained effort to collect data over long time periods. This topic is taken up in some detail in the following section, which focuses on research design.

4 RESEARCH DESIGN

The previous section illustrates that many different methods are available for poverty impact assessment—ranging from statistical approaches to qualitative methods. Because of the complexity of poverty impacts and the multiplicity of the underlying pathways, all of these methods may be needed in different settings. We do not argue here that one method is correct and others are not, nor do we propose to dictate the choice of methods to researchers aiming to document poverty impacts. Instead, we believe that a portfolio of research will be needed to establish clear links from agricultural research to poverty impacts, with different methods and techniques used to document different components of the causal chain.

What is clear is that the selection of methods is in fact a secondary concern for impact assessment. The primary concern is research design, rather than research method. Research design in this case includes the selection ab initio of geographic areas for research, the conceptualization of time frames over which impact is likely to take place, the extent of adoption and diffusion that are anticipated, the likely pathways to impact, and the potential magnitudes of the effects to be measured or documented. Too often, existing studies have been carried out belatedly and with fragmentary data that cannot hope to provide a coherent picture of impacts.

This paper argues that for high-quality studies of poverty impacts, the research design effort must begin before the technology is actually introduced to farmers—and in fact the research design must be a joint effort of the scientists introducing innovations and the social scientists who are charged with the impact assessment. Research design must be forward looking, even when the impact assessment itself is backward looking. This in turn requires a degree of institutional cooperation and planning that has been rare in recent experience. The following sections provide some general thoughts on research design. For each method, there are research design issues to be addressed; conversely, for different research questions, there are methods that are more or less suitable. The point is to develop a research design that is internally consistent, so that the central research questions can plausibly be answered with the selected methods.

4.1 INGREDIENTS OF RESEARCH DESIGN

The key logical elements of a research design for poverty impact assessment are given by the need to measure poverty outcomes in a setting where a new technology has been introduced and to offer a plausibly valid counterfactual for the outcomes that would have been expected in the absence of the innovation. The statistical issues described above ("the impact evaluation problem") are specific examples, formulated in the language of statistics, of the broader problem posed by the need for meaningful and persuasive counterfactuals.

For a poverty impact assessment for an agricultural technology, a checklist for research design might include the following elements:

- An assessment of the geographic area in which the technology is likely to diffuse, perhaps a recommendation domain or an agro-ecological zone to which the technology is well suited.
- An assessment of the time horizon over which diffusion is expected to unfold, with a corresponding assessment of the time until which the impact of the technology is likely to be realized.
- An understanding of the key pathways for impact and hence the populations that are likely to experience the most significant impacts.
- Development of an appropriate counterfactual scenario: what scenario would likely unfold in the absence of the new technology?
- Considered estimates of the magnitudes of the impacts likely to be experienced by different populations.
- Clearly defined and validated metrics that can be used to monitor uptake of the technology and its impacts.
- Ideally, some means of assessing attribution for the innovation.

Given the time lags involved in diffusion and impact, this research design is intrinsically forward looking, often over a period of a decade or more. It need not begin at the moment at which a technology is first introduced to farmers, but the research design will be cleanest when it can be conducted relatively early in the diffusion of a new technology.

4.2 COUNTERFACTUALS AND THE RANDOMIZATION BENCHMARK

For precisely the reasons that some agricultural research is conducted in the public sector, the technologies that emerge cannot often be introduced in a random fashion to farmers or communities. The underlying problem is that technologies such as seeds or management practices are freely diffusible from farmer to farmer, making it nearly impossible over any extended period of time to maintain an uncontaminated "control" group for any successful technology "treatment" that is introduced. This relatively free movement of technologies across farmers explains the reluctance of the private sector to invest in research in these areas; in many settings, there is limited scope for firms to recover a share of the benefits from the new technology, and hence limited scope for them to recoup the cost of research investments. The public sector role in research (including that of CGIAR) is driven by precisely this market failure.

This difficulty of maintaining distinct treatment and control groups may also make randomization difficult or impossible as an evaluation strategy over moderate to long periods of time. Nevertheless, randomization offers a useful heuristic benchmark for research design. Our search for a valid counterfactual can generally be satisfied in a research design with appropriate randomization. But in the absence of randomization, it is still sensible to seek a research design that allows for a useful and potentially persuasive counterfactual.

At the very least, the need for a counterfactual will often dictate monitoring outcomes in some area or within some population that is *not* directly affected by the new technology. Perhaps this means carrying out surveys in an adjacent area that is unsuitable for the new technology, or perhaps it implies monitoring the well-being of households that are not involved in agriculture. Perhaps there is a plausibly similar population in another country or region where the research is not yet suitable.

The point is not that the research design must claim a truly random assignment of subjects to treatment and control; it is that there must be serious thought given to the establishment of a plausible counterfactual. When this is done *ex ante*, it is far more credible than when it is done *ex post*. At present, researchers typically do little more than compare adopters with non-adopters, but for the reasons already discussed, these comparisons seldom offer valid counterfactuals.

Some possible approaches that fall short of full randomization include using phased rollouts of new technologies, randomizing the location of intensive extension services or pilot testing of new technologies, or using the purchase price of inputs associated with the technology (e.g., new seeds) as a method of varying the timing and intensity of technology diffusion.

Sometimes there is no convincing control population or geography; in this case, the only alternative may be to use a model-based counterfactual. But, in this situation, at least the model can be calibrated to the baseline time period so that its validity can then be evaluated *ex post*: the calibrated model, with the new technology added to it, can then be compared with the realizations of the data, making it possible to evaluate the model's external validity. This is perhaps preferable to a model that is constructed *ex post* and forced to fit the data. In the latter case, it is always difficult to tell how much the model can be trusted as a representation of the actual economy.

4.3 AVOIDING SAMPLE SELECTION BIAS AND "STRATEGIC SITE" BIAS

A recurring problem in the research design for impact assessment is sample selection bias. Far too many CGIAR impact studies are guilty of focusing on samples of farmers where adoption is high. In the extreme case, this can actually involve selection on the dependent variable—normally regarded as a serious error in sampling. More frequently, the problem results from selection on an independent variable that is highly correlated with the outcome variable of interest. For instance, researchers may limit the sample for an adoption study to a set of districts where adoption levels are high. This approach skews the sample in such a way that the results have little external validity. In most cases, the study sample should be chosen to be convincingly representative of some larger population. This need not be a nationally or even regionally representative population, but the sample must at least be selected on the basis of some exogenous characteristic that is not intimately related to the outcome variable of interest. Choosing survey sites that are unusually interesting—rather than those that are statistically representative—is a formula for arriving at results that are not meaningful.

This point is highly relevant in the context of current discussions in CGIAR about so-called sentinel sites, where data can be collected simultaneously for the evaluation of multiple research programs. The problem with this approach is that any locations that are of interest to multiple research programs are ipso facto atypical. These are likely to be locations with many different commodities (which implies that they are agroecologically atypical) and with fairly high connectivity and good infrastructure (or else they would not be chosen). This more or less automatically implies that what happens in these locations is not representative of what is happening in other locations.

An implication is that data collection may often involve sites and locations that researchers view as unfavorable or undesirable.

4.4 COMMITMENT TO APPROPRIATE DATA COLLECTION

Different innovations affect poverty through different pathways, and impact studies need to document achievement toward poverty reduction using measures that are consistent with the specific pathways and methods to be used. The appropriate outcome measures need to be defined at the outset, so that meaningful baseline data can be collected. If an innovation is intended to reduce poverty, then household measures of yield or crop production cannot provide the needed information; instead, it will be necessary to demonstrate a change over time in one of the generally accepted measures of poverty[3].

[3] This paper will not venture into the crowded literature on poverty measures; this is a rich area of development economics. There are absolute and relative measures of poverty based on income, expenditures, and assets (among other things). There are multidimensional measures of poverty and direct measures based on anthropometry and health status. These are all defensible ways of measuring poverty.

In much of the literature that we reviewed for this paper, impacts were limited to the yield or production of a particular crop or production system targeted for the innovation. But farmers reallocate inputs and effort in complicated ways, so observing the full effects of a technology "treatment" normally requires monitoring across a range of farm and household outcomes. For example, the introduction of a high-yielding variety of a staple food may induce an increase in farm labor, pulling it away from other income-generating activities; the net effect on farm income may be smaller than the gross effect on the production value of the targeted crop.

In practice, this means that poverty-oriented impact assessments often need to collect data on the entire farm-household portfolio of economic activities. This in turn may require buy-in from researchers, who may struggle to understand why an impact assessment of a new maize variety requires collecting detailed data on cassava production—or, for that matter, on the profits of household nonfarm enterprises. A related implication is that data collection for poverty impact assessment is likely to be more detailed and therefore more costly than data collection for a simple adoption study. Research budgets need to factor in the costs of this kind of data collection—including baseline studies that will eventually be necessary for assessing changes and impact.

Another dimension in which data collection will be challenging is the length of time over which impacts are expected to take place. For many of the innovations produced by CGIAR, diffusion and impact may stretch over decades; it would certainly not be uncommon for an innovation to take 10–15 years to demonstrate measurable poverty impacts. To document and assess impacts over this kind of time frame requires sustained attention, commitment, and support. In some cases, the necessary data will actually take the form of long-term panels; in other cases, repeated cross-sections may provide sufficient information. Either way, data collection must cover sufficiently long periods of time for impacts to be realized.

Issues of statistical power are also important in research design. In many cases, the likely poverty impacts will be modest. Given the highly stochastic nature of agricultural production, as well as the year-on-year variability in prices, business cycles, and other essentially random shocks, it is likely to take large samples and long periods before poverty impacts can reliably be distinguished from statistical noise. It will not in general be possible to observe poverty impacts with any confidence over short periods (say, less than five years), and even over much longer time periods, it may prove extremely difficult to find statistically meaningful poverty impacts. It is critical that research design should draw on power calculations that guide both the sample size and the study duration needed to deliver significant results, given the likely extent of fluctuations in the data. This is true regardless of the eventual method to be used for analyzing the results; even qualitative studies of poverty impacts need to be designed in such a way that they can plausibly separate statistical signal from noise.

4.5 SUMMARY

Effective research design for poverty impact assessment will require extensive forward planning, lengthy and detailed data collection, and clear development of counterfactual scenarios. It should be clear that this research will also require significant investments of resources; these are not studies that can easily or cavalierly be carried out through quick-and-dirty cross-section surveys. With appropriate planning and investment, however, it should be possible to set up research designs that will eventually allow for credible estimation of poverty impacts.

5 CONCLUSION

This paper makes several arguments that can be summarized as follows:

First, poverty impact assessment is feasible but difficult. There are multiple pathways through which agricultural research can lead to changes in the prevalence and severity of poverty. Not every innovation will lead to significant poverty impacts, and poverty impact is a complex outcome; this means that the relationship will be difficult to tease out from data. Nevertheless, with sufficiently large samples and sufficiently long time periods, the statistical relationship should emerge from the data. The same argument applies to qualitative studies; no single study will necessarily yield a convincing picture of poverty impacts, but with enough data, it should be possible to find a compelling relationship between research and poverty impacts.

Second, the main challenge to poverty impact assessment is not one of research methods. There are many methods that can effectively and legitimately be brought to bear on the problem. Almost any method, applied rigorously, can offer useful evidence on poverty impacts. There is no reason to insist on a single methodology or a single approach.

Third, we argue that much more important—and more challenging—than the specific quantitative or qualitative method is the design of the research. Too often in the past, impact studies have been undertaken on the basis of poor research designs, leading to situations where the studies are trying to ask questions that the data and methods cannot honestly address. Good research design requires far more planning, often quite early in the process of technology adoption and diffusion. Issues such as sample design, measurement, statistical power, and the development of appropriate counterfactuals are critical; no statistical method can compensate for the failure to think through these issues clearly.

Fourth, effective research design in the context of poverty impact assessment for CGIAR will require close collaboration between social scientists and the researchers who are responsible for the development and diffusion of new technologies. Researchers who are committed to poverty impact assessment must understand that the success of the impact assessment will depend in part on their willingness to cooperate with respect to the rollout of new technologies across time and space, among other things.

The challenges of poverty impact assessment are substantial; it is difficult enough to demonstrate the impact of new technologies on productivity, which is a much more direct causal relationship. But with sufficient attention to research design and planning, and sufficient commitment of resources, there is the potential to assemble sets of results that can provide compelling evidence on the poverty impacts of CGIAR research. The magnitude of the challenge should not be underestimated. But the possibilities for learning are also great, and so, correspondingly, is the potential to improve the system so that it can more effectively reach the poor.

6 REFERENCES

Abro, Z. A., Alemu, B. A., & Hanjra, M. A. 2014. Policies for agricultural productivity growth and poverty reduction in rural Ethiopia. *World Development*, 59, 461–474.

Adato, M., & Meinzen-Dick, R., eds. 2007. *Agricultural research, livelihoods, and poverty: Studies of economic and social impacts in six countries*. Baltimore, MD: Johns Hopkins University Press.

Alene, A. D., & Coulibaly, O. 2009. The impact of agricultural research on productivity and poverty in sub-Saharan Africa. *Food Policy*, 34, 198–209.

Anríquez, G., & López, R. 2007. Agricultural growth and poverty in an archetypical middle income country: Chile 1987–2003. *Agricultural Economics,* 36, 191–202.

Antle, J. M., Murshed-E-Jahan, K., & Crissman, C. C. 2013. Moving along the impact pathway: Improved methods for estimating technology adoption and multi-dimensional impact: The case of integrated aquaculture-agriculture in Bangladesh. A report submitted to the Standing Panel on Impact Assessment (SPIA) of CGIAR Independent Science and Partnership Council (ISPC).

Asfaw, S., Kassie, M., Simtowe, F., & Lipper, L. 2012. Poverty reduction effects of agricultural technology adoption: A micro-evidence from rural Tanzania. *Journal of Development Studies*, 48, 1288–1305.

Becerril, J., & Abdulai, A. 2010. The impact of improved maize varieties on poverty in Mexico: A propensity score-matching approach. *World Development*, 38, 1024–1035.

Behagel, L., Giginoux, J., Macours, K., Mangheni, M., Beraho, M., & Kamugisha, R. 2014. Making extension work in Uganda. Nairobi: World Agroforestry Centre (ICRAF).

Bezu, S., Kassie, G. T., Shiferaw, B., & Ricker-Gilbert, J. 2014. Impact of improved maize adoption on welfare of farm households in Malawi: A panel data analysis. *World Development*, 59, 120–131.

Burke, M., & Miguel, E. 2014. Innovative finance for technology adoption in Western Kenya. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab (J-PAL). https://www.povertyactionlab.org/evaluation/innovative-finance-technology-adoption-western-kenya.

Bustos, P., Caprettini, B., & Ponticelli, J. 2013. Agricultural productivity and structural transformation. Evidence from Brazil. Chicago Booth Research Paper No.14-07; Fama-Miller Working Paper.

Carter, M. R., Laajaj, R., & Yang, D. 2013. The impact of voucher coupons on the uptake of fertilizer and improved seeds: Evidence from a randomized trial in Mozambique. *American Journal of Agricultural Economics*, 95, 1345–1351.

Carter, M., Lybbert, T. J., & Mathenge, M. K. 2014. Evaluating the socio-economic impacts of Western Seed's Hybrid Maize Program: Baseline report. Nairobi: Tegemeo Institute, Egerton University.

Christiaensen, L., Demery, L., & Kuhl, J. 2011. The (evolving) role of agriculture in poverty reduction: An empirical perspective. *Journal of Development Economics*, 96, 239–254.

Cole, S., & Fernando, A. N. 2012. The value of advice: Evidence from mobile phone-based agricultural extension. Harvard Business School Working Paper, No. 13–047, November 2012.

Cunguara, B., & Darnhofer, I. 2011. Assessing the impact of improved agricultural technologies on household income in rural Mozambique. *Food Policy*, 36, 378–390.

Dar, M. H., Emerick, K., de Janvry, A., & Sadoulet, E. 2014. The impact of seed fairs on the diffusion of new varieties. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab (J-PAL), Agricultural Technology Adoption Initiative (ATAI).

Datt, G., & Ravallion, M. 1998a. Farm productivity and rural poverty in India. *Journal of Development Studies*, 34, 62–85.

Datt, G., & Ravallion, M. 1998b. Why have some Indian states done better than others at reducing rural poverty? *Economica*, 65, 17–38.

David, C. C., & Otsuka, K. 1994. *Modern rice technology and income distribution in Asia*. Los Baños, Philippines: International Rice Research Institute.

Davis, K., Nkonya, E., Kato, E., Mekonnen, D. A., Odendo, M., Miiro, R. & Nkuba, J. 2012. Impact of farmer field schools on agricultural productivity and poverty in East Africa. *World Development*, 40, 402–413.

de Janvry, A., & Sadoulet, E. 2009. Agricultural growth and poverty reduction: Additional evidence. *World Bank Research Observer*, 25, 1–20.

de Janvry, A., Dustan, A., & Sadoulet, E. 2011. Recent advances in impact analysis methods for ex-post impact assessments of agricultural technology: Options for the CGIAR. Report prepared for the workshop "Increasing the rigor of ex-post impact assessment of agricultural research," organized by the CGIAR Standing Panel on Impact Assessment (SPIA), October 2, 2010, Berkeley, CA.

De Laat, J., Jack, W., Kremer, M., & Suri, T. 2014. Encouraging adoption of rainwater harvesting tanks through collateralized loans in Kenya. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab (J-PAL), Agricultural Technology Adoption Initiative (ATAI). https://www.povertyactionlab.org/evaluation/encouraging-adoption-rainwater-harvesting-tanks-through-collateralized-loans-kenya.

Dercon, S., Gilligan, D. O., Hoddinott, J., & Woldehanna, T. 2009. The impact of agricultural extension and roads on poverty and consumption growth in fifteen Ethiopian villages. *American Journal of Agricultural Economics*, 91, 1007–1021.

Dercon, S., & Gollin, D. 2014. Agriculture in African development: A review of theories and strategies. *Annual Review of Resource Economics*, 6, 471–492.

Diao, X., & Thurlow, J. 2014. *The impact of agricultural technology adoption on growth and poverty reduction in Ethiopia, 2002–2013*. Washington, DC: International Food Policy Research Institute (IFPRI).

Dillon, A. 2011a. Do differences in the scale of irrigation projects generate different impacts on poverty and production? *Journal of Agricultural Economics*, 62, 474–492.

Dillon, A. 2011b. The effect of irrigation on poverty reduction, asset accumulation, and informal insurance: Evidence from northern Mali. *World Development*, 39, 2165–2175.

Dillon, A., Sharma, M., & Zhang, X. 2011. Estimating the impact of rural investments in Nepal. *Food Policy*, 36, 250–258.

Dorosh, P. A., & Mellor, J. W. 2013. Why agriculture remains a viable means of poverty reduction in Sub-Saharan Africa: The case of Ethiopia. *Development Policy Review*, 31, 419–441.

Duflo, E., Glennerster, R., & Kremer, M. 2007. Using randomization in development economics research: A toolkit. *Handbook of Development Economics*, 4, 3895–3962.

Duflo, E., Keniston, D., & Suri, T. 2014. Promoting agricultural technology adoption in Rwanda. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab (J-PAL).

Duflo, E., Kremer, M., & Robinson, J. 2008. How high are rates of return to fertilizer? Evidence from field experiments in Kenya. *American Economic Review*, 98, 482–488.

Duflo, E., Kremer, M., & Robinson, J. 2009. *Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya*. Cambridge, MA: National Bureau of Economic Research.

Emerick, K., de Janvry, A., Sadoulet, E., & Dar, M. H. 2014. Risk and the modernization of agriculture. Working paper, Department of Economics, Tufts University.

Evenson, R. E., & Gollin, D. 2003. Assessing the impact of the Green Revolution, 1960 to 2000. *Science*, 300, 758–762.

Fan, S., Hazell, P., & Haque, T. 2000. Targeting public investments by agro-ecological zone to achieve growth and poverty alleviation goals in rural India. *Food Policy*, 25, 411–428.

Fan, S., & Pardey, P. G. 1997. Research, productivity, and output growth in Chinese agriculture. *Journal of Development Economics*, 53, 115–137.

Filipski, M., Manning, D., Taylor, J. E., Diao, X., & Pradesha, A. 2013. *Evaluating the local economy-wide impacts of irrigation projects*. IFPRI Discussion Paper 01247. Washington, DC: International Food Policy Research Institute.

Freebairn, D. K. 1995. Did the Green Revolution concentrate incomes? A quantitative study of research reports. *World Development*, 23, 265–279.

Gajwani, K., Kanbur, R., & Zhang, X. 2007. *Comparing the evolution of spatial inequality in China and India: A fifty-year perspective*. DSGD Discussion Paper No. 44. Washington, DC: International Food Policy Research Institute.

Glennerster, R., & Suri, T. 2014a. The impact of new rice varieties on health in Sierra Leone. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab (J-PAL), Agricultural Technology Adoption Initiative (ATAI).

Glennerster, R., & Suri, T. 2014b. *Promoting adoption of new rice varieties: Addressing the costs of early adoption in Sierra Leone*. Agricultural Technology Adoption Initiative.https://www.povertyactionlab.org/evaluation/promoting-adoption-new-rice-varieties-addressing-costs-early-adoption-sierra-leone.

González-Flores, M., Bravo-Ureta, B. E., Solís, D., & Winters, P. 2014. The impact of high value markets on smallholder productivity in the Ecuadorean Sierra: A stochastic production frontier approach correcting for selectivity bias. *Food Policy*, 44, 237–247.

Hamazakaza, P., Smale, M., & Kasalu, H. 2013. The impact of hybrid maize on smallholder livelihoods in Zambia: Findings of a household survey in Katete, Mkushi, and Sinazongwe Districts. Lusaka, Zambia: Indaba Agricultural Policy Research Institute.

Hasan, R., & Quibria, M. G. 2004. Industry matters for poverty: A critique of agricultural fundamentalism. *Kyklos*, 57, 253–264.

Hazell, P. B. 2010. An assessment of the impact of agricultural research in South Asia since the Green Revolution. *Handbook of Agricultural Economics*, 4, 3469–3530.

Hazell, P. B., & Ramasamy, C. 1991. *The Green Revolution reconsidered: The impact of high-yielding rice varieties in South India*. Baltimore, MD: Johns Hopkins University Press.

Hayami, Y. & Kikuchi, M. 2000. A Rice Village Saga: Three Decades of Green Revolution in the Philippines. London, UK: Macmillan Press.

Holden, S. 2013. Input subsidies and demand for improved maize: Relative prices and household heterogeneity matter! As, Norway: Centre for Land Tenure Studies, Norwegian University of Life Sciences.

Holden, S., & Mangisoni, J. 2013. Input subsidies and improved maize varieties in Malawi: What can we learn from the impacts in a drought year? As, Norway: Centre for Land Tenure Studies, Norwegian University of Life Sciences.

Hope, R. 2007. Evaluating social impacts of watershed development in India. *World Development*, 35, 1436–1449.

Hornbeck, R., & Keskin, P. 2014. The historically evolving impact of the Ogallala Aquifer: Agricultural adaptation to groundwater and drought. *American Economic Journal: Applied Economics*, 6, 190–219.

Ivanic, M., & Martin, W. 2014. *Sectoral productivity growth and poverty reduction: National and global impacts*. Washington, DC: World Bank.

Jewitt, S., & Baker, K. 2007. The Green Revolution re-assessed: Insider perspectives on agrarian change in Bulandshahr district, Western Uttar Pradesh, India. *Geoforum*, 38, 73–89.

Kabunga, N. S., Dubois, T., & Qaim, M. 2014. Impact of tissue culture banana technology on farm household income and food security in Kenya. *Food Policy*, 45, 25–34.

Kanbur, R., & Shaffer, P. 2007. Epistemology, normative theory and poverty analysis: Implications for Q-squared in practice. *World Development*, 35, 183–196.

Karlan, D., & Udry, C. 2014. Evaluating combined farm productivity interventions in Ghana. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab (J-PAL), Agricultural Technology Adoption Initiative (ATAI). https://www.povertyactionlab.org/fr/evaluation/examining-underinvestment-agriculture-returns-capital-and-insurance-among-farmers-ghana

Kassie, M., Pender, J., Yesuf, M., Kohlin, G., Bluffstone, R., & Mulugeta, E. 2008. Estimating returns to soil conservation adoption in the northern Ethiopian highlands. *Agricultural Economics*, 38, 213–232.

Kassie, M., Shiferaw, B., & Muricho, G. 2010. *Adoption and impact of improved groundnut varieties on rural poverty: Evidence from rural Uganda*. Environment for Development Discussion Paper 10-11. Washington, DC: Resources for the Future, Environment for Development.

Kathage, J., & Qaim, M. 2012. Economic impacts and impact dynamics of Bt (Bacillus thuringiensis) cotton in India. *Proceedings of the National Academy of Sciences*, 109, 11652–11656.

Kerr, J. M., & Kolavalli, S. 1999. Impact of agricultural research on poverty alleviation: Conceptual framework with illustrations from the literature. EPTD Discussion Paper No. 56. Washington, DC: International Food Policy Research Institute.

Kijima, Y., Otsuka, K., & Sserunkuuma, D. 2008. Assessing the impact of NERICA on income and poverty in central and western Uganda. *Agricultural Economics*, 38, 327–337.

Kleemans, M., & Sadoulet, E. 2014. Impact of drought-tolerant risk-reducing rice on yield and farmer welfare in India. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab (J-PAL), Agricultural Technology Adoption Initiative (ATAI). https://www.povertyactionlab.org/evaluation/impact-drought-tolerant-risk-reducing-rice-yield-and-farmer-welfare-india.

Lanjouw, P. & Stern, N. 1998. Economic development in Palanpur over five decades. Oxford, UK: Oxford University Press. ISBN 9780198288329

Larochelle, C., Alwang, J., Norton, G. W., Katungi, E., & Labarta, R. A. 2013. Ex-post impact of improved bean varieties on poverty and food security in Uganda and Rwanda. Cali, Colombia: International Centre for Tropical Agriculture (CIAT).

Ligon, E., & Sadoulet, E. 2008. Estimating the effects of aggregate agricultural growth on the distribution of expenditures. Background paper for the *World Development Report 2008*. Washington, DC: World Bank.

Lipton, M., & Longhurst, R. 1989. *New seeds and poor people*, Baltimore, MD: Johns Hopkins University Press.

Loayza, N. V., & Raddatz, C. 2010. The composition of growth matters for poverty alleviation. *Journal of Development Economics*, 93, 137–151.

Mahajan, A., Seira, E., Corral, C., & Giné, X. 2014. A multiple interventions approach to increasing technology adoption: Evidence from Mexico. https://basis.ucdavis.edu/project/multiple-interventions-approach-increasing-technology-adoption-view-towards-scaling

Maheshwari, A. 1998. Green Revolution, market access of small farmers and stagnation of cereals' yield in Karnataka. *Indian Journal of Agricultural Economics*, 53, 27–40.

Majid, N. 2004. Reaching Millennium Goals: How well does agricultural productivity growth reduce poverty? Geneva: International Labour Organization.

Maredia, M., Byerlee, D., & Anderson, J. R. 2000. Ex post evaluation of economic impacts of agricultural research programs: A tour of good practice. Revised version of a paper presented to the Workshop on "The Future of Impact Assessment in CGIAR: Needs, Constraints, and Options", May 3-6, 2000, Rome: Standing Panel on Impact Assessment (SPIA) of the Technical Advisory Committee.

Mathenge, M. K., Smale, M., & Olwande, J. 2014. The impacts of hybrid maize seed on the welfare of farming households in Kenya. *Food Policy*, 44, 262–271.

Matuschke, I., Mishra, R. R., & Qaim, M. 2007. Adoption and impact of hybrid wheat in India. *World Development*, 35, 1422–1435.

McKenzie, D. 2012. Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, 99, 210–221.

Mendola, M. 2007. Agricultural technology adoption and poverty reduction: A propensity-score matching analysis for rural Bangladesh. *Food Policy*, 32, 372–393.

Minten, B., & Barrett, C. B. 2008. Agricultural technology, productivity, and poverty in Madagascar. *World Development*, 36, 797–822.

Murray, M. P. 2006. Avoiding invalid instruments and coping with weak instruments. *Journal of Economic Perspectives*, 20, 111–132.

Ndjeunga, J., Umar, J., Ahmed, B., Aba, A., Ibro, A., Abdoulaye, A., & Gwadi, K. 2011. Adoption and impacts of modern sorghum and pearl millet varieties in northern Nigeria. Patancheru, India: International Crops Research Institute for the Semi-Arid Tropics.

Noltze, M., Schwarze, S., & Qaim, M. 2013. Impacts of natural resource management technologies on agricultural yield and household income: The system of rice intensification in Timor Leste. *Ecological Economics*, 85, 59–68.

Pamuk, H., Bulte, E., Adekunle, A., & Diagne, A. 2014a. Decentralised innovation systems and poverty reduction: Experimental evidence from Central Africa. *European Review of Agricultural Economics*, 42, 99–127.

Pamuk, H., Bulte, E., & Adekunle, A. A. 2014b. Do decentralized innovation systems promote agricultural technology adoption? Experimental evidence from Africa. *Food Policy*, 44, 227–236.

Pande, R., & Islam, H. 2014. Reducing imbalanced fertilizer use in Bangladesh. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab (J-PAL), Agricultural Technology Adoption Initiative (ATAI). https://www.povertyactionlab.org/evaluation/reducing-imbalanced-fertilizer-use-bangladesh.

Prahladachar, M. 1983. Income distribution effects of the Green Revolution in India: A review of empirical evidence. *World Development*, 11, 927–944.

Raitzer, D. A., Sparks, A. H., Huelgas, Z., Maligalig, R., Balangue, Z., Launio, C., Daradjat, A., & Ahmed, H. U. 2013. *Is rice improvement still making a difference? Assessing the economic, poverty and food security impacts of rice varieties released from 1989 to 2009 in Bangladesh, Indonesia and the Philippines*. Report submitted to the Standing Panel on Impact Assessment (SPIA), CGIAR Independent Science and Partnership Council (ISPC).

Ravallion, M., & Chen, S. 2007. China's (uneven) progress against poverty. *Journal of Development Economics*, 82, 1–42.

Rosegrant, M. W., & Hazell, P. B. 2000. *Transforming the rural Asian economy: The unfinished revolution*. Oxford, UK: Oxford University Press.

Sanglestsawai, S., Rejesus, R. M., & Yorobe, J. M. 2014. Do lower yielding farmers benefit from Bt corn? Evidence from instrumental variable quantile regressions. *Food Policy*, 44, 285–296.

Shaffer, P. 2013. Ten years of "Q-squared": Implications for understanding and explaining poverty. *World Development*, 45, 269–285.

Shiferaw, B., Kassie, M., Jaleta, M., & Yirga, C. 2014. Adoption of improved wheat varieties and impacts on household food security in Ethiopia. *Food Policy*, 44, 272–284.

Smale, M., & Mason, N. 2014. Hybrid seed and the economic well-being of smallholder maize farmers in Zambia. *Journal of Development Studies*, 50, 680–695.

Sovey, A. J., & Green, D. P. 2011. Instrumental variables estimation in political science: A readers' guide. *American Journal of Political Science*, 55, 188–200.

Subramanian, A., & Qaim, M. 2010. The impact of Bt cotton on poor households in rural India. *Journal of Development Studies*, 46, 295–311.

Taylor, J. E., Filipski, M., Thome, K., Kagin, J., & Davis, B. 2013. Agricultural spillover effects of cash transfers: What does LEWIE have to say? *American Journal of Agricultural Economics*, 95, 1338–1344.

Taylor, J. E., & Filipski, M. J. 2014. *Beyond experiments in development economics: Local economy-wide impact evaluation*. Oxford, UK: Oxford University Press.

Thapa, G., Otsuka, K., & Barker, R. 1992. Effect of modern rice varieties and irrigation on household income distribution in Nepalese villages. *Agricultural Economics*, 7, 245–265.

Thirtle, C., Irz, X., Lin, L., McKenzie-Hill, V., & Wiggins, S. 2001. *Relationship between changes in agricultural productivity and the incidence of poverty in developing countries*. Report commissioned by the Department for International Development, London.

Thirtle, C., Lin, L., & Piesse, J. 2003. The impact of research-led agricultural productivity growth on poverty reduction in Africa, Asia and Latin America. *World Development*, 31, 1959–1975.

White, H. 2002. Combining quantitative and qualitative approaches in poverty analysis. *World Development*, 30, 511–522.

Zeng, D., Alwang, J., Norton, G. W., Shiferaw, B., Jaleta, M., & Yirga, C. 2013. Ex-post impacts of improved maize varieties on poverty in rural Ethiopia. 2013 Annual Meeting, August 4-6, 2013, Washington, D.C. 149823, Agricultural and Applied Economics Association.

APPENDIX 1: LITERATURE REVIEW ON STUDY RESULTS

This appendix summarizes the key results of the papers examined in the literature review. We have attempted to summarize the main findings concerning the relationships between agricultural research, productivity change, and impacts on poverty. This is not intended to be a comprehensive evidence review, but it provides some insight on the main findings of the literature. The first three papers were prepared in the context of the Organizational Change Program for CGIAR funded by the Ford Foundation in the late 1990s.

1. OBSERVATIONAL MICRO STUDIES OF DIRECT FARM-LEVEL IMPACTS

Most of the observational studies find that the adoption of modern varieties is associated with outcomes that can potentially be linked to poverty:[4]

- An increase in yields, except possibly under drought conditions (Holden and Mangisoni, 2013).
- An increase in total household income in many cases; exceptions include the case of sorghum in northern Nigeria (Ndjeunga et al., 2011), a watershed development program in India (Hope, 2007), a farmer field school in Uganda (Davis et al., 2012), and four improved technologies in a drought year in Mozambique (Cunguara and Darnhofer, 2011).
- An increase in per capita expenditure, except for pearl millet and sorghum in northern Nigeria (Ndjeunga et al., 2011).
- A reduction in poverty for the few studies that reported results; an important exception was a study of the impact of extension in Nepal (Dillon et al., 2011).
- Improved food security.

In terms of magnitude of impacts, estimates are difficult to compare across studies. The locations as well as the types of crops and technologies investigated are quite different. The variables employed to measure both the adoption of technologies and the outcomes vary considerably between studies (e.g., within the studies focusing on maize, four consider adoption a binary variable at the household level, one measures adoption at the plot level, two treat adoption as a continuous variable, and one defines adoption as the area planted to hybrids).

Estimates of the positive impact of improved varieties on crop income are available for India, where hybrid wheat has an overall net income advantage of US$ 39 per acre, representing 14 percent of mean per capita total expenditure among adopters (and 20 percent of expenditure for non-adopters), based on research by Matuschke et al. (2007). Improved groundnut in Uganda was seen to result in a US$ 159–180/ha increase in net groundnut crop income (about 35 percent), and improved pearl millet in Nigeria was associated with a 48–92 percent increase in per capita value of crop production (Ndjeunga et al., 2011). A study of Bt cotton suggested that it increased profits per acre by 50 percent over conventional cotton (Kathage and

4 Unless stated otherwise, all results reported are average treatment effects on the treated (ATT).

Qaim, 2012). Soil conservation practices resulted in an 18–25 percent increase in value of crop production in low-rainfall areas of Ethiopia (although there was no significant difference for high-rainfall areas). Davis et al. (2012) estimated that farmer field schools in Kenya and Tanzania were associated with 21 percent and 61 percent increases in agricultural incomes respectively.

Panel-based studies on improved maize adoption find increases of total household income of 7 percent in Kenya (Mathenge et al., 2014) and 18 percent[5] in Zambia (Smale and Mason, 2014), as well as an increase of 0.26 percent per 1 percent of area under improved maize for Malawi (Bezu et al., 2014). A much larger impact—a 64 percent increase in total income—is estimated by Hamazakaza et al. (2013) for hybrid maize in Zambia; this study, however, is relatively small (300 households) and based on only a cross-section. For rice, cross-sectional studies find that high-yielding variety (HYV) rice adoption in Bangladesh is associated with a 30 percent higher total household income (Mendola, 2007), whereas the system of rice intensification (SRI) was found to produce an increase of only 2.34 percent (Noltze et al., 2013). A large impact is found for banana tissue culture technology, with an 89–116 percent (US$ 500–662) rise in annual farm household income in Central/Eastern Kenya (Kabunga et al., 2014). Again this is a relatively small cross-sectional study of 385 households, even though the authors argue this is in line with estimates from *ex ante* impact assessments.

These increases in income do not always translate into comparably large increases in consumption expenditure. For instance, for Bt cotton adoption Kathage and Qaim (2012) based on panel data find no impact at early stages of adoption (2002–2004) but increases of 18 percent (US$ 321) four years later (2006–2008). Improved maize in Mexico is associated with an increase in per capita consumption of 4.6–4.9 percent for adopters (Becerril and Abdulai, 2010). Much higher estimates on the same outcome are found in a cross-sectional study of chickpea in Ethiopia (24.6 percent), and even higher (103 percent) for pigeonpea in Tanzania (Asfaw et al., 2012). Results for the impact of extension and irrigation are mixed. In studies using panel data, the findings have been quite uneven: Dillon (2011b) from northern Mali finds a rise of total household consumption of 25–28 percent associated with irrigation and Dercon et al. (2009) finds receiving at least one extension visit in Ethiopia increases consumption by 7.1 percentage points, whereas Dillon et al. (2011) finds neither of the two has an impact on consumption in Nepal.

Compared with the impacts on income and consumption, the poverty reduction estimates are mostly moderate: 1 kg of maize hybrid seeds planted is associated with 0.172 percentage-point reduction in poverty severity in Zambia (Smale and Mason, 2014) and with a 0.29 percentage-point reduction in poverty depth in Kenya (Mathenge et al., 2014). Adopters of hybrid maize have a 31 percent lower probability of being poor in Mexico (Becerril and Abdulai, 2010). Adopters of HYV rice have a 14 percent lower probability of being poor in Bangladesh (Mendola, 2007). Receiving at least one extension visit is associated with a 9.1 percentage-point reduction in poverty headcount in Ethiopia (Dercon et al., 2009).

An alternative to studies that actually attempt to show reductions in poverty are those that look at differential impacts on households within different income groups. Several studies find that the benefits of new technologies are at least as strong for poorer farm households as for wealthier ones. The evidence is limited, however: hybrid maize adoption in Malawi increases the income of the poorest 30 percent of households, whereas it has no significant impact on the richest 30 percent. A similar study on pearl millet adoption in northern Nigeria shows higher yields and crop value for the poorest 25 percent of farm households only (Ndjeunga et al., 2011). In Timor Leste, both poor and nonpoor households benefit from SRI adoption in a similar magnitude (Noltze et al., 2013).

How impacts vary with farm size is less clear-cut.

[5] 1 kg of seeds was associated with an US$ 8 increase in total household income, which is equivalent to about 0.5 percent of the sample mean income in the two periods. The mean amount of hybrid seeds planted was 36.7 kg, which would result in an income increase of 18 percent.

Three studies find technology benefits to be inversely related to farm size, at least in proportional terms. This is true in the case of hybrid wheat in India where small farms (<5 acres) gain 5 percent more per acre than medium-size farms (5–25 acres) and 40 percent more than large farms (>25 acres) (Matuschke et al., 2007). Note, however, that these results suggest that benefits are higher for small farms in *relative* terms—but in absolute terms, the larger farms probably increase their production more. In Timor Leste, the relative income gain of intensified rice production for small farms (<2 ha) appears to be higher than that of large farms (4.8 percent versus 0.07 percent income gain) (Noltze et al., 2013). The same is true for hybrid maize adoption in Mexico, where per capita consumption rises more strongly for small farms (<5 hectares) than for large farms (Becerril and Abdulai, 2010). However, several papers also find benefits to be concentrated among middle-sized farms. Adopting improved chickpea in Ethiopia and improved pigeonpea in Tanzania seems to be associated with increases in per capita expenditure that are largest for the second and third quintiles of the farm-size distribution (Asfaw et al., 2012). Likewise, improved groundnut in Uganda results in income gains that are most pronounced for the second and third farm-size quintiles (Kassie et al., 2010). In Bangladesh, the benefits of improved rice adoption appear to increase with land owned: income effects for near-landless are only half of the mean, whereas for medium-large farms they are 30–60 percent higher than the average (Mendola, 2007). Finally, Davis et al. (2012) find barely any significant impact of Farmer Field Schools (FFS) on small and large farms but increases of between 20 and 100 percent for almost all countries and indicators for medium-size farms. These results could be a reflection of the diversity of technologies and conditions covered by the different studies. To some extent, they may also be explained by the different understandings of large and small farms. (Some studies use an absolute size threshold; others compare size quintiles, implicitly defining large and small in relative terms.)

As expected, more-educated individuals are more likely not only to adopt but also to benefit more strongly from adoption. Adoption of improved pea varieties in Tanzania and Ethiopia has the strongest positive impacts on per capita expenditure of households in the third and fourth educational quintiles (Asfaw et al., 2012). Income benefits from adoption of improved groundnut are highest for households in the fourth educational quintile (even though benefits are evenly distributed among the other quintiles), and a significant poverty reduction is only found for those in the third and fifth educational quintile (Kassie et al., 2010). Further, in Dercon et al. (2009) extension visits appear to have higher positive impacts for households with a younger and more educated head. In contrast to this, FFS seem to mainly benefit households with no education (increases of 40–250 percent in crop and livestock productivity and income) (Davis et al., 2012).

Only two studies differentiate impacts by gender. One finds that improved maize adoption results in a stronger increase of maize own-consumption in female-headed households, but no further gender-related differences in impact on income or assets are observed (Bezu et al., 2014). The other study finds that FFS have a much larger benefit for female-headed households for three outcome indicators (crop and livestock productivity and income) and all three countries (Kenya, Tanzania, Uganda) (Davis et al., 2012).

The findings in this literature thus suggest that the benefits of improved varieties or technology are not limited to large farms with wealthier, male-headed, and more-educated farm households but that they may extend to relatively poor smallholder farmers. However, these regressions on subgroups are all likely to suffer from problems of the reduction of sample size (particularly in studies with only 300–400 households), which calls for some caution regarding the reliability of results.

2. OBSERVATIONAL MICRO STUDIES OF INDIRECT IMPACTS

Even though the improved varieties result in considerable increases in yields (40–80 percent), estimates of the impact on household net income and poverty are rather moderate. Improved beans in Rwanda and Uganda under a small open economy (SOE) assumption result in an increase of annual farm household income of US$ 73.49 and US$ 62.32 (PPP), respectively, and a reduction in the poverty headcount of 0.4 and 0.1 percentage points (Larochelle et al., 2013). Improved maize in Ethiopia has a slightly larger impact: in the SOE case, the headcount is reduced by 0.8–1.2 percentage points; for the closed economy this effect is slightly weaker, with a 0.6–0.9 percentage-point headcount reduction (Zeng et al., 2013). In both cases, the effect on poverty depth and severity is even lower. Raitzer et al. (2013) takes a slightly different approach and only evaluates the share of the additional surplus generated by the improved variety allocated to the poor. Results reveal that under the assumption of a positive shutdown supply price an important share of benefits accrue to the poor: about 40 percent to those living on less than PPP$ 2 a day in Indonesia and Philippines and about 20 percent to those living on less than PPP$ 1.25 a day; the shares are even larger in Bangladesh, at 66 percent and 50 percent. This allocation is not affected by replacing the positive shutdown supply with a constant elasticity (CE) supply function. However, under a CE supply function the additional surplus generated by the improved varieties largely diminishes (or even entirely vanishes). In conclusion, while a considerable share of benefits accrues to the poor, the overall impact on poverty is rather moderate and sensitive to assumptions about the tradability of the agricultural good and the nature of the supply function.

In contrast, Subramanian and Qaim (2010) use a micro social accounting matrix (SAM) multiplier model and find that even though growing additional Bt cotton in an Indian village would considerably raise aggregate household income, this increase would barely be captured by the very poor (<US$ 1.15/day) or the landless. It would, however, benefit vulnerable households (between US$ 1.15 and 2.3 /day) to some extent and would notably increase aggregate returns to hired female and non-agricultural labor. The validity of these results may be limited by the strong assumptions underlying this type of model (fixed prices, perfectly elastic supply of factors and resources, SAM specific to village even though it may to some extent be typical for the semi-arid tropics).

These assumptions particularly clash with findings from Minten and Barrett (2008) who, in a meso-level study, analyze the relationship between yields and prices or wages in commune-level data. They find that local prices respond quite strongly to local rice yields and that communes with higher yields or cash crops have higher real wages. However, both the data source (obtained in part from focus group discussions) and the econometric approach (2SLS using share of land under improved irrigation and proportion of population belonging to forest ethnic group as IV) of this paper are problematic.

3. EXPERIMENTS AND QUASI-EXPERIMENTS

There are not yet many results from the randomized trials in our literature review, as many are still at early stages. Some specific findings include the following:

- Traits of an improved variety that reduce risk have an impact on farmer decision making (area planted, inputs allocated [labor, fertilizer], credit and saving decisions) (Emerick et al., 2014).
- Improved varieties may require specific training (households that adopted NERICA-3 but received no training had lower yields than control households). The timing of variety yields may be key (potentially improved health outcomes among households that grow NERICA-3 likely to be working through earlier maturity of the crop) (Glennerster and Suri, 2014b). One-time input voucher subsidies can have persistent learning effects beyond the crop or the individual they are targeted to, but only if a strong exposure to the technology has occurred before (Carter et al., 2013; Duflo et al., 2009).
- Behavioral aspects may be relevant for adop-

tion of a technology (Duflo et al., 2009), but results from experimental data may still need to make certain assumptions that crucially affect results (Duflo et al., 2008).
- Farmers use mobile phone–based extension services, but results are moderate and more pronounced among more-educated farmers (Cole and Fernando, 2012).
- Innovation platforms appear to promote adoption of new technologies among all types of households in eight sub-Saharan African countries, however extent to which technology adoption is promoted is limited (mainly involving crop management techniques), and they may take some time to establish (Pamuk et al., 2014b).
- Innovation platforms in three sub-Saharan African countries may reduce poverty more strongly than conventional extension services, but the only poverty outcome available is an aggregate village poverty measure, which is the result of subjective "focus group discussions" within the village (Pamuk et al., 2014a).

The 11 currently ongoing trials identified are running mainly in sub-Saharan Africa: Bangladesh, Ghana, India (two), Kenya (three), Mexico, Rwanda, Sierra Leone, and Uganda. Several trials center on modern rice varieties (drought-tolerant rice, NERICA plus low-cost training, Swarna-Sub1), and two trials look at hybrid maize. Four studies investigate the interdependencies between the adoption of modern technology and access to credit or insurance. Finally, two studies investigate the impact of extension/ agronomy training. The outcomes assessed are similar to those of observational studies and completed trials: adoption (six) and diffusion (three) of new varieties, the impact of adoption on farm decision-making, farm yields and production (six), children's health, and expenditures on health and education. Three trials only examine the impact of adoption on household income or welfare. One study explicitly investigates the impact of modern varieties on girls' time allocation, girls' school enrollment, and the intrahousehold allocation of resources.

4. MACROECONOMETRIC STUDIES

Our literature review included four papers that explicitly examined the impact of agricultural research on poverty. These papers conclude that for sub-Saharan Africa and Asia, research-induced technical change accounts for an important share of agricultural output growth. (Estimates of the total productivity elasticity with respect to agricultural research fall in the range of 0.34–0.38). The time periods under scrutiny vary, but in most cases cover the 1980s and 1990s. Evidence from India suggests that the elasticity is considerably affected by the agro-ecological zone and the availability of complementary inputs. Thus, Fan et al. (2000) find that high-yielding varieties increased productivity in irrigated but not in rainfed areas. For China, Fan and Pardey (1997) further find that the marginal returns from research increase through time, with an annual growth of the elasticity coefficient of about 4.1 percent. The elasticity of poverty (defined as per capita income/ expenditure less than US$ 1/day) with respect to agricultural research varies considerably across the studies, but it is difficult to tell how much of this reflects measurement and estimation differences and how much is attributable to the difference in continents and time periods. For sub-Saharan Africa in the period 1980–2003, Alene and Coulibaly (2009) estimate an elasticity of -0.22. But, Thirtle et al. (2003) obtain a much larger value for sub-Saharan Africa for 1985–1990 (-0.717) and also find a larger value for Asia (-0.48) and a notably lower figure for the Americas (-0.153).

With respect to the impact of growth in the agricultural sector (measured either as growth of sectoral labor productivity or as value added), studies mostly find that it is has larger poverty-reducing impacts than growth originating in other sectors (see, for example, Ligon and Sadoulet, 2008; Majid, 2004; Christiaensen et al., 2011; Thirtle et al., 2003; Ravallion and Chen, 2007; Loayza and Raddatz, 2010). This is particularly valid for sub-Saharan Africa and South Asia and for the extremely poor (<US$ 1/day). Many of these papers find that growth in the nonagricultural sector does not significantly contribute to poverty reduction at all

(Ligon and Sadoulet, 2008; Majid, 2004; Thirtle et al., 2003). The strong poverty-reducing impact is also corroborated by evidence from Loayza and Raddatz (2010), who identify a sector's intensity of unskilled labor (high in agriculture) as a main determinant of its poverty-reducing capacity.

Not all the results in this literature support the idea that poverty impacts from agricultural productivity or agricultural growth are greater than for other sectors. de Janvry and Sadoulet (2009), Christiaensen et al. (2011), and Hasan and Quibria (2004) find that growth in the secondary and tertiary sectors is more important for poverty reduction among the less poor (those living on between US$ 1 and US$ 2 per day) and also in Latin America and East Asia. Authors suggest that the smaller share of the agricultural sector in some of these countries, as well as pronounced land inequality, could explain these findings (Christiaensen et al., 2011; de Janvry and Sadoulet, 2009).

5. MACRO MODELS

To the extent that macro models have addressed the poverty impacts of research, they have generally supported the findings of the macroeconometric studies. For instance, based on a 31-country CGE model, Ivanic and Martin (2014) find that among the three sectors productivity increases in agriculture have the largest potential to reduce poverty. The magnitude of the effect is positively related to the size of the productivity improvement but not affected by the number of countries adopting the improvement. (Note again, however, that the model is silent about the cost or difficulty of achieving productivity increases in the different sectors.) Their findings also support the notion that the main poverty effects operate through lower living costs, which also benefit poor urban households. They suggest that the effect on the latter will be more pronounced in case of relatively widespread productivity improvements.

Similarly, Dorosh and Mellor (2013), based on findings from a three-sector growth model calibrated to data from Ethiopia, conclude that high agricultural growth, particularly among small commercial farmers, can considerably increase employment (including in the rural nonfarm sector) and through these multiplier effects, accelerate poverty reduction and promote urbanization. Findings from Anríquez and López (2007), to some extent, contradict the rather pessimistic econometric results for Latin America in Thirtle et al. (2003). Using a structural model to compensate for shortcomings of household data (only a pseudo-panel is available), they find that agricultural growth in Chile outperforms growth in other sectors when it comes to poverty reduction. The elasticities of poverty they obtain are comparable to those found for sub-Saharan African countries, which they attribute to high labor market fluidity.

In contrast to this, findings from Diao and Thurlow (2014) are far less optimistic. Combining a CGE model with microeconometric evidence, they evaluate to what extent yield improvements have contributed to poverty reduction in Ethiopia. They concentrate on the four main cereal crops and yield increases attributable to improved seeds, increased use of fertilizer, and extension. Based on simulations, they conclude that increased technology use and extension have contributed to only 3.3 percent of the rise in agricultural GDP and 3.5 percent of the poverty headcount reduction observed during 2001–2013. In addition, even full adoption of these technologies would result in only a 12.3 percent increase in agricultural GDP and a reduction in current poverty levels of one-third.

6. QUALITATIVE METHODS

It is useful to briefly highlight some of the key insights of this literature and to discuss methodological issues for this kind of qualitative research. The Green Revolution importantly contributed to poverty reduction in Asia, mainly through lower food prices (Hazell, 2010). It was not able, however, to entirely eliminate poverty and malnutrition, in part due to the very high population growth that simultaneously occurred (Rosegrant and Hazell, 2000). Further, the Green Revolution largely failed to improve existing regional inequalities, particularly those related to agro-ecological conditions. In India, the fact that some of the poorest regions (relying on rainfed agriculture only) barely participated in the Green Revolution even resulted in widening income disparities (Gajwani et al., 2007; Prahladachar, 1983). Interregional seasonal migration of agricultural workers may have been able to mitigate these effects to some degree (David and Otsuka, 1994). Some of the disadvantaged regions were also able to catch up once they received irrigation technologies. Evidence on direct farm-level effects is inconclusive. Even though early studies suggested that benefits were concentrated heavily on larger and better-endowed farms (Freebairn, 1995), more recent evidence emphasizes the ambiguity and context dependence of the impacts. In some cases, small-scale farmers experienced the largest gains, particularly in the long run (Hazell and Ramasamy, 1991; Thapa et al., 1992; Maheshwari, 1998; Jewitt and Baker, 2007; Lipton and Longhurst, 1989). One important lesson seems to be that a technology's benefits for large versus small farms depends far less on the characteristics of the technology than on the prevailing socioeconomic conditions: a less unequal land and income distribution, and well-functioning institutions and markets increase the probability of observing more equitable social impacts of a technology (Kerr and Kolavalli, 1999; Hazell, 2010; Adato and Meinzen-Dick, 2007; Freebairn, 1995). Freebairn (1995) identifies a further interesting aspect: a study's conclusions about whether the Green Revolution increased or reduced inequality depends on its structural and methodological approach: the most favorable social outcomes are found in micro-based case studies, whereas the worst outcomes are reported by macro-based essays.

7. UNDER-RESEARCHED TOPICS

The literature review reveals a number of significant gaps in our knowledge of poverty impacts. These gaps include an almost complete lack of evidence on many of the indirect effects of agricultural research through effects on structural changes in the economies. We lack evidence concerning the impacts of agricultural research on distributional issues, such as the distribution of landholdings and wages. Other topics that have been inadequately documented are the following:

- Many of the indirect effects of research through pathways other than prices and quantities.
- Impact on labor markets and wages (for which our review found only the results in Subramanian and Qaim 2010).
- Impact on intrahousehold dynamics: women's and children's time allocation and bargaining position.
- Impact on migration decisions and outcomes.
- Impact of technology adoption on food prices and price variation under different local infrastructures.

Independent
Science and
Partnership
Council

CGIAR Independent Science & Partnership Council (ISPC) Secretariat
c/o FAO, Viale delle Terme di Caracalla 00153 Rome, Italy
t: +39 06 570 52103
http://ispc.cgiar.org