What do we really know about the impacts of improved grain legumes and dryland cereals?

A critical review of 18 impact studies

Erik Katovich, Andrew Feist, Karl Hughes, Kai Mausch

World Agroforestry

What do we really know about the impacts of improved grain legumes and dryland cereals?

A critical review of 18 impact studies

Erik Katovich, Andrew Feist, Karl Hughes, Kai Mausch

About the authors

Erik Katovich is a PhD student in Agricultural and Applied Economics at the University of Wisconsin, Madison. He previously completed a Fulbright Research Fellowship in Development Economics (2016) at the State University of Campinas in Brazil and holds a Bachelor of Science in Economics (2015) from the University of Minnesota. Email: ekatovich@wisc.edu

Andrew Feist is a PhD student in Agricultural and Applied Economics at the University of Wisconsin, Madison. He holds a Bachelor of Science in Mathematics from Rensselaer Polytechnic Institute (2015). Email: awfeist@wisc.edu

Karl Hughes is an Impact Evaluation and Monitoring and Evaluation specialist who heads ICRAF's Impact Unit. He obtained his PhD, focusing on impact evaluation methodology, from the London School of Hygiene and Tropical Medicine, University of London (2012). He also holds a Master's in Environmental Studies (1998) from York University, Canada and a Bachelor's Degree in Anthropology and Philosophy (1994), from the University of British Columbia, Canada. Email: k.hughes@cgiar.org

Kai Mausch has been working on rural development for more than 10 years, exploring solutions to rural poverty from both the agricultural and the non-agricultural perspectives. Before joining ICRAF, he worked for the International Crops and Research Institute for the Semi-Arid Tropics (ICRISAT) from 2010 to 2018 where he planned, coordinated and implemented economic projects and programme components at the regional and global levels. Topics ranged from adoption, dissemination and impact evaluation to value chain analysis and improved targeting of agricultural interventions in rural areas. Kai received his MSc while working with the International Centre of Insect Physiology and Ecology (ICIPE) in Nairobi and received his PhD in Economics from the School of Economics and Management, Leibniz University of Hanover, Germany (2009). Email: k.mausch@cgiar.org

Acronyms

ATE   average treatment effect

DREAM   Dynamic Research Evaluation for Management

FPVS   farmer participatory variety selection

GLDC   grain legume and dryland cereal

ICIPE   International Centre of Insect Physiology and Ecology

ICRISAT   International Crops Research Institute for the Semi-Arid Tropics

IV   instrumental variables

OLS   ordinary least squares

PSM   propensity score matching

R&D   research and development

RCT   randomized control trial

ROI   return on research investment

Abstract

Improved grain legume and dryland cereal (GLDC) varieties hold potential to intensify smallholder agriculture and improve livelihoods in semi-arid regions of sub-Saharan Africa and South Asia. To assess the empirical evidence base for these potential benefits, we review 18 GLDC impact studies and identify gaps in current knowledge on GLDC impacts. Results from this synthesis reveal that all five reasonably well-identified adoption studies estimate significant, positive effects of improved GLDC adoption on yields, profits, or household welfare. Another well-identified study focuses on nutritional impacts of improved GLDC consumption and measures positive effects on iron-deficiency in school children. Macro-level welfare estimates based on economic surplus models (eight of the 18 studies) are largely invalidated because of their dependence on poorly-identified household-level impact estimates. Four additional studies rely on correlations and expert interviews. Overall, impact studies focus on chickpea and groundnut, as opposed to other GLDC crops. Studies are geographically concentrated in Ethiopia, India, and Tanzania, and are heavily focused on estimating economic impacts, with few studies assessing potential environmental, nutritional or social impacts. Recommendations are offered to improve methodological approaches in future impact assessments of GLDC crops.

Table of contents

1. Introduction

Smallholder farmers in semi-arid regions of sub-Saharan Africa and South Asia face severe challenges, including drought, soil degradation, malnutrition, and competition for land. Many households in these areas depend on staple crops such as maize, rice and wheat, which often fail during times of drought and – unless complemented by other foods – do not provide essential proteins and micronutrients (Huho and Mugalavai, 2010). Households may also grow unimproved grain legume and dryland cereal (GLDC) varieties such as pigeonpea, cowpea, groundnut, millet or sorghum, which have the potential for improving nutrition and soil fertility, but frequently suffer from diseases and low yields (Shiferaw, 2008).

Population growth in semi-arid regions contributes to unsustainable intensification and extensification of land use, as well as high rates of poverty and food insecurity. Poverty rates in these areas are significantly higher than continental averages, ranging from 49% in semi-arid zones of South Asia to 83% in the sub-humid zones of East Africa (Jayne et al., 2014; Gumma, 2017). The adverse effects of climate change, felt most drastically in fragile semi-arid agro-ecologies, have the potential to deepen these problems and complicate development efforts. The scope for production gains through expansion into untouched arable lands is limited or negative due to the exhaustion of frontier spaces, desertification, urbanization and other factors (De Janvry, 2001).

Thus, efforts to combat poverty, food insecurity, malnutrition, and environmental degradation in semi-arid zones of sub-Saharan Africa and South Asia must include a focus on the sustainable intensification of smallholder production (Simtowe, 2012). Improved GLDC crops have potential to contribute significantly to this effort. GLDC crops provide essential proteins, oils and micronutrients to smallholder households that often cannot afford animal products (Ghosh et al., 2007). Legume-cereal intercropping and rotation systems also improve soil quality through legume nitrogen fixation and reduce the risk of crop failure through diversification. Improved varieties offer greater yields, drought and disease resistance, and valuable market opportunities (Valbuena, 2012).

Despite these potential benefits, GLDC crops have long been seen as "subsistence crops", and national and international research centers have historically neglected research, development and promotion of improved GLDC varieties in favor of staple cereals like maize (Shiferaw, 2008). Recently however, agricultural research institutions have re-prioritized crop and diet diversification as important contributors to food security (Mayes et al., 2012). For example, the Consultative Group on International Agricultural Research (CGIAR) initiated the *Grain Legumes and Dryland Cereals Agri-Food Systems* programme in 2018 with a focus on increasing the "productivity, profitability, resilience and marketability of critical and nutritious grain legume and cereal crops grown within the semi-arid and sub-humid dryland agro-ecologies of sub-Saharan Africa and South Asia (CGIAR, 2018)." The programme seeks to promote intensified crop research

focused on GLDC crops, as well as the adoption at scale of improved varieties and associated agronomic practices such as intercropping.

While the positive contributions of improved GLDC crops to soil fertility, nutrition, and yields are well-established in experimental settings, their impacts at scale, in farmers' fields, and over the long-term are less clear. Many farmers fail to adopt appropriate management practices along with improved GLDC crops, attenuating the benefits of nitrogen fixation and pest and disease management. Others opt to sell improved GLDC crops and consume staples or local varieties, thus missing out on potential nutritional gains. Still others recycle certified seeds, diminishing the yield and gains in drought and disease resistance achieved by agricultural research stations under ideal conditions. Furthermore, effecting change in semi-arid agro-ecological systems is complicated by geographical dispersion and remote populations, weak infrastructure and institutions, and severe biophysical constraints (Dercon and Gollin, 2014). For these and myriad other reasons related to the inherent complexity of smallholder agricultural production, the ability and cost-effectiveness of improved GLDC crops to combat poverty, malnutrition, and environmental degradation is not self-evident. It is therefore essential to rigorously evaluate the impacts of the promotion and adoption of improved GLDC varieties on the ground.

Impact evaluations involve rigorous assessment of changes associated with specific policy interventions or practices. Ideally, evaluations identify the causal effects of an intervention on "treated" units relative to a control group. In practice however, data and resource constraints often limit evaluators' ability to identify causal effects. Methods of impact assessment range from randomized controlled trials and quasi-experimental methods such as difference-in-difference and regression discontinuity design, to observational approaches such as instrumental variables (IVs) and propensity score matching (PSM), correlational studies, expert interviews, and opinion surveys (Gertler, 2016). See Appendix 1 for a review of these methods and their application to GLDC adoption.

Numerous studies have sought to estimate the impacts of improved GLDC crops on farm households' welfare, incomes, consumption, and poverty. Further studies have sought to estimate returns on research investment (ROI) in GLDC crop breeding and promotion. While nearly all studies have estimated positive welfare gains and competitive ROI from improved GLDC crops, the data and methodologies employed in these studies range widely in quality and appropriateness, raising doubts regarding the magnitude and nature of their findings. These doubts, considered in light of growing investments in GLDC crop improvement and promotion, highlight the urgent need to establish a reliable and coherent evidence base in this area.

The intent of this study is to fill this knowledge gap and drive impact assessment methodologies in productive new directions by systematically compiling and reviewing recent, relevant evidence of impacts resulting from the adoption of improved GLDC crops in sub-Saharan Africa and South Asia. The 18 studies included in this review were obtained through keyword searches in the archives of the International Crops Research Institute for the Semi-arid Tropics (ICRISAT), Google Scholar, and AgEcon Search. For each study, we critically evaluate the quality and

representativeness of the data, the quality and appropriateness of the research methodology, and the results. Based on these assessments, we assess the degree to which each study's findings can be generalized to broader contexts. We then draw conclusions regarding the effectiveness of past crop research and prospects for future gains from research or scaling. Our goal is to synthesize what is currently known, identify gaps in knowledge and methodological approaches, and generate recommendations for good practices for future impact evaluations of GLDC crop research and adoption.

The study is presented as follows: section 2 critically reviews the existing literature on impact studies for improved GLDC crops; section 3 discusses the current state of the evidence on GLDC impacts; and section 4 draws conclusions and offers recommendations for future impact evaluations. Detailed summaries and critiques of each study are presented in appendices.

2. Literature review

This section reviews the recent literature on impact studies of improved GLDC varieties in sub-Saharan Africa and South Asia. Appendix 2 contains a full description of the criteria used to compile the studies as well as summaries and critiques of each study. Appendix 3 contains summary tables of impact studies and their findings. The studies reviewed in this section range in methodology from randomized experimental designs to IVs, matching, and panel data approaches, to macro-level welfare estimates and correlations. Studies cover most major GLDC crops and a selection of countries. Refer to section 4 for a discussion of the current state of GLDC impact evidence as a whole.

Experimental evidence on improved GLDC impacts is limited. The only farm-level experimental evidence comes from Bulte et al. (2014), who administer unblinded and double-blinded randomized trials of improved cowpea adoption in Tanzania. Findings from this study indicate that, in the "naïve" unblinded randomized controlled trial (RCT), improved cowpea adoption is associated with a 27% increase in yields relative to local varieties, whereas in the double-blind RCT, two-thirds of this measured impact is a placebo effect from treatment-induced effort. These findings generate uncertainty regarding the contributions of improved crop genetics relative to behavioral responses, leaving the door open for further farm-level experiments. The only other experimental evidence in this literature is provided by Finkelstein et al. (2015), who administer a double-blind RCT of improved pearl millet consumption in an Indian boarding school. Results from the trial indicate that improved pearl millet consumption is associated with significant reductions in iron deficiency among schoolchildren, highlighting the positive contributions of an improved GLDC crop to nutrition.

While experimental evidence on improved GLDC impacts is lacking, a number of observational studies employ panel data, IVs, and propensity-score matching methods to identify causal effects. Michler et al. (2018) offer a rejoinder to Bulte et al. by highlighting the importance of measuring household-level – rather than field-level – outcomes. The authors measure the effects of improved chickpea adoption on household yields and profits, and take advantage of panel data to control for differences across households that do not vary over time, and localized rainfall data to control for household-level shocks. They find that adoption has negligible effects on yield but significant positive effects (23-28%) on profits, and conclude that since households optimize at the household level, this is the correct level at which to conduct impact analysis.

Verkaart et al. (2017) and Asfaw et al. (2012) provide further evidence of the positive effects of improved GLDC adoption on household incomes, welfare, and poverty reduction. Verkaart et al. draw on panel data from Ethiopia and apply a two-stage IV procedure to control for the endogeneity of chickpea adoption decisions. The authors estimate that a 10% increase in area planted with improved chickpea is associated with a 12.6% increase in income per capita and a 3.9% reduction in probability of being below the US$2/day poverty line. Asfaw et al. use "contact with extension agents" and "participation in farmer participatory variety selection (FPVS) trials" to instrument for adoption of improved chickpeas (in Ethiopia) and pigeonpea (in Tanzania), and

estimate that adoption of improved varieties is associated with positive welfare effects of between 24.6% (Ethiopia) and 103% (Tanzania). While results from both studies corroborate each other, the strength of Asfaw et al.'s exclusion restriction[1] is highly questionable, leading to the exclusion of their findings from the body of "better-identified evidence" presented in Part 4.

Finally, impact studies of improved groundnut adoption by Njeunga et al. (2013) and Simtowe et al. (2012) – focused on Nigeria and Malawi respectively – apply PSM to reduce endogeneity from selection bias in crop adoption. Both studies find evidence that adoption of improved groundnut varieties has a significant positive effect on household consumption (and on poverty reduction in Malawi). Nonetheless, the inability of PSM techniques to control for unobservable heterogeneity suggests that these results should be treated with caution.

The largest subset of the GDLC impact literature focuses on country-level estimates of welfare and ROI in crop improvement (Tsusaka et al., 2016; Kumara Charyulu, 2016a; Kumara Charyulu, 2016b; Bantilan, 2014; Gierend, 2014; Macharia, 2012; Shiferaw, 2008). These studies typically proceed by estimating farm-level gains from adoption and the extent of adoption, and then extrapolate from these values to arrive at a net welfare impact. All of the studies in this category rely on poorly-identified farm-level effects and more or less unrepresentative samples, thereby invalidating many of their quantitative findings. Tsusaka et al. (2016) offer an excellent summary of the evolution of the groundnut sector and agricultural policy in Malawi since the 1980s, but confront substantial data limitations when attempting to estimate the country-level welfare impact of groundnut research. Their use of 1977-1982 groundnut yields as a counterfactual to current yields incorrectly attributes all post-1982 improvements in groundnut productivity to crop research. Most other studies in this category measure impact using simple yield comparisons between adopters and non-adopters, thus failing to control for selection-into-adoption and biasing their estimates upwards if more productive farmers are more likely to adopt.

All of the country-level welfare studies cited in the previous paragraph estimate large, positive welfare impacts from improved GLDC variety adoption, as well as positive ROI from crop research. This regularity in findings is compatible with pervasive selection bias in these studies' farm-level impact estimates, which may bias estimates upwards. Furthermore, given that the calibration of welfare models requires a series of assumptions, there is substantial scope for conscious or unconscious researcher bias, especially considering that most studies are written by researchers evaluating the success of their own organization's work. It is therefore possible that the positive welfare estimates from these studies are spurious. On the other hand, estimates of welfare gains from crop improvement align with findings from the better-identified studies reviewed above, as well as with substantial theoretical and case-study evidence highlighting the importance of agricultural research (Pardey et al., 2006).

---

[1] The exclusion restriction states that the instrument in question should only affect the outcome variable indirectly by influencing treatment status (Angrist and Pischke, 2008).

In principle, country-level estimates of ROI relative to alternative investment options (agricultural extension, downstream value-addition, manufacturing, etc.) are of great interest, especially considering the divergence of opinion regarding the most effective sectoral priorities for agrarian economies (World Development Report, 2008; Dercon and Gollin, 2014). Future studies of this kind should take advantage of existing representative surveys and better-identified impact studies to arrive at more reliable and useful welfare estimates.

Finally, a number of GLDC impact studies rely on correlations or modeling methods to assess the impacts of improved GLDC varieties. These studies offer valuable context or historical analysis, but cannot distinguish well-identified impacts. Kaliba et al. (2017) draw on survey data from Tanzania to generate distributions of yields and socioeconomic indicators for local versus improved sorghum varieties, and conclude that the distribution of net returns from improved varieties stochastically dominates that of local varieties. Considering the prevalence of risk aversion and precarity in smallholder agriculture, Kaliba et al.'s careful consideration of distributions – rather than point estimates – is a welcome contribution. Nonetheless, the observed mean and variance shifts in net returns are endogenous correlations, since it is unlikely that farmers' choices about crops and agronomic practices are independent of other factors determining outcomes.

In India, Kumara Charyulu et al. (2014) measure positive district-level correlations between rates of improved sorghum adoption and sorghum yields, and Karunakaran et al. (2013) measure a positive association between FPVS trials and improved variety adoption. Both studies fail to account for selection into adoption, likely biasing their estimates upwards. Finally, Nwankwo et al. (2010) employ a pure modeling approach to estimate the cost-savings from soybean nitrogen fixation in sub-Saharan Africa. The authors calculate that total soybean cultivation across 19 African countries resulted in a fertilizer-cost-equivalent savings of US$203 million over the 2002-2004 period. This study constitutes the only available assessment of environmental impacts of GLDC adoption in the literature, and notably does not employ survey data. Clearly, there is room for future studies to make important contributions regarding the environmental impacts of improved GDLC crops and associated management practices.

3. Discussion

Impacts from improved GLDC varieties flow from smallholder farmers' decisions to adopt these new crops and associated management practices. These impacts may be immediate or long-term, and may also be dynamic, changing in nature and magnitude as the ratio of adopters to non-adopters evolves (see for example, the discussion of the "technology treadmill" in Sunding and Zilberman, 2001). Impacts from the adoption of improved GLDC crops may potentially be felt across five areas:

Microeconomic: changes in household income, consumption, poverty, and market orientation;
Macroeconomic: changes in state- or country-level welfare, agricultural productivity, institutional capacity, and agro-economic structure;
Environmental: changes in soil, water, and air quality, biodiversity, carbon sequestration, and climate resilience;
Nutrition and food security: changes in diet, nutrition, health and human development outcomes, and hunger; and
Social: changes in bargaining power and social and gender relations within households and communities, farmers' aspirations, and behavior (including migration decisions).

Investment in the development and promotion of improved GLDC varieties for smallholder agriculture in sub-Saharan Africa and South Asia is based on the conclusion that the impacts of these investments across the thematic areas above are positive, cost-effective, and sustainable. Indeed, the theoretical impact pathways underlying this conclusion are compelling (CGIAR, 2018). Nonetheless, it is important that policy be evidence-driven as well as theory-driven. The objective of this paper is therefore to compile and critique the existing evidence on impacts from improved GLDC crops, and to identify gaps in our current knowledge of GLDC impacts.

3.1 Current state of the evidence

Figure 1 presents the number of GLDC impact studies by crop, country, area of impact, and methodology. Of the impact studies reviewed, five focus on chickpea, four on groundnut, four on sorghum, two on pigeonpea, two on pearl millet, and one each on finger millet, cowpea, and soybean. Better-identified causal evidence of positive farm-level impacts is strongest for improved chickpea and groundnut (Michler et al., 2018; Verkaart et al., 2017; Ndjeunga et al., 2013; Simtowe et al., 2012)). Furthermore, Bulte et al. (2014) provides experimental evidence of small but positive yield-impacts for improved cowpea; Finkelstein et al. (2015) present experimental evidence of positive health and nutrition impacts for iron-fortified pearl millet. Importantly, all of these studies should be interpreted in light of their limited external validity and potential identification problems.

Well-identified evidence on impacts from improved sorghum, pigeon pea, pearl millet, finger millet, cowpea, and soybean is lacking. Studies bringing representative datasets and appropriate impact-evaluation methods to bare on these crops would offer a substantial contribution to the evidence base.

GLDC impact studies were conducted in Ethiopia (four), India (six), Tanzania (four), Nigeria (one), Malawi (two), and Uganda (one). This collection of countries captures important regional diversity (covering East, West, and Southern Africa, and South Asia). Nevertheless, sub-Saharan Africa and South Asia are characterized by substantial heterogeneity in agro-ecological conditions and economic structures, limiting the external validity of impact studies and requiring locally specific crop breeding and agronomic recommendations. For this reason, impact studies expanding the focus beyond Ethiopia, India, and Tanzania, or focusing on less-studied regions within these countries, would constitute a significant contribution.

FIGURE 1. NUMBER OF GLDC IMPACT STUDIES BY (CLOCKWISE FROM TOP LEFT) CROP, COUNTRY, IMPACT THEME, AND METHODOLOGY[2]

---
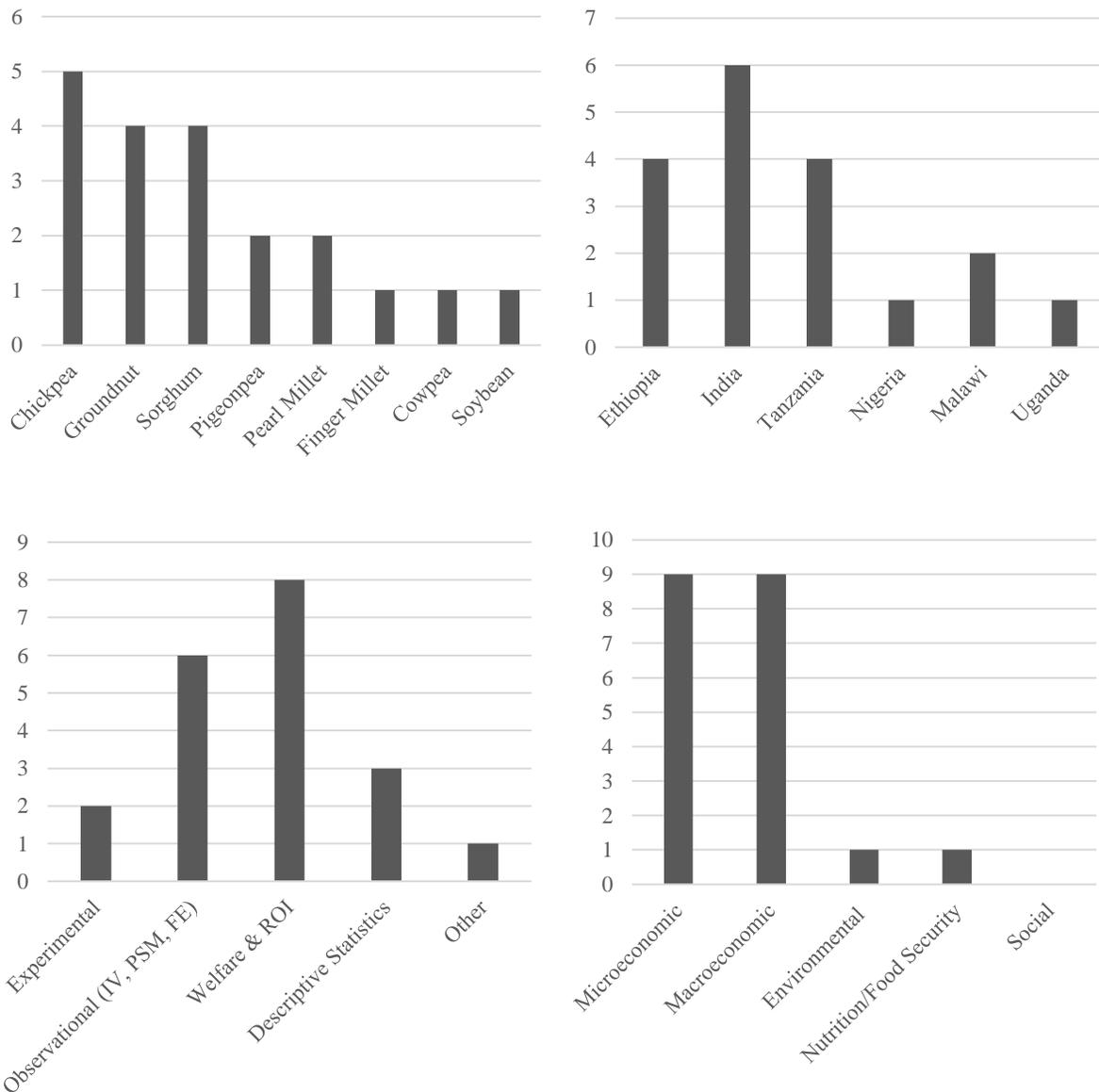
[2] The number of studies reported varies slightly among the different figures because some studies incorporate multiple countries, crops, methods, or impact areas.

As illustrated in Figure 1, nearly all relevant impact studies of improved GLDC crops have focused on micro and macroeconomic impacts, which are relatively easier to measure than environmental, nutritional, or social impacts. Largely for this reason, impact studies for these latter areas are lacking. Only one study (Finkelstein et al., 2015) is directly focused on assessing the health and nutrition impacts of a GLDC crop (iron-fortified pearl millet). And only one study (Nwankwo et al., 2010) is focused on measuring environmental benefits (cost savings from soybean nitrogen fixation). While there are attempts to outline the interplay between cropping decisions and social structures, no impact studies focused on measuring impacts in these spheres (e.g. changes in household gender relations, community dynamics, or migration) were found despite the substantial focus on these factors in project proposals. Clearly, the evaluation of health, food security, environmental, and social impacts is a high priority going forward.

Of the studies reviewed, two were experimental, six used observational methods such as IV, PSM, and fixed effects, eight used macro-level economic surplus models, and four used descriptive statistics or other methods. Given their comparative advantage in identifying causal effects, further experimental evaluations of improved GLDC crops would be important contributions to the evidence base. Furthermore, the majority of economic surplus approaches (i.e. state- or country-level welfare and ROI studies) are invalidated by poorly-identified farm-level impact estimates. Moving forward, it is important for researchers to couple economic surplus modeling with better-identified farm-level effects of adoption. See the Conclusion section for further recommendations regarding methodology.

3.2 What we know
The better-identified causal evidence available suggests that, in certain contexts, the adoption of improved GLDC crops improves household-level yields, incomes, and consumption, and the consumption of nutritious GLDC food products reduces iron deficiency (see Figure 2).[3]

As argued in Michler et al. (2018), evidence of positive GLDC impacts may be even stronger if studies focus on measuring economic returns rather than yields. On the other hand, most crop-specific impact studies sample from regions best suited for the production of that crop, suggesting that gains may diminish if scaling moves beyond these ideal agro-ecologies.

The available estimates of state- and country-level welfare impacts and ROI are largely invalidated by their failure to identify farm-level causal effects of adoption. Interestingly, all of these studies find positive welfare impacts and ROI. While this may be a result of researchers' confirmation bias, it may also offer suggestive evidence (aligning with theory and intuition) that investments in crop improvement and agricultural development improve welfare and yield competitive returns. Improvements in future welfare and ROI analyses may be realized by

---

[3] We opt here for the term "better-identified," rather than "well-identified" to acknowledge the presence of substantial measurement and estimation problems within some of these studies. Nonetheless, the "better-identified" results presented in Figure 2 seem sufficiently robust to allow meaningful interpretation.

incorporating better-identified farm-level impact estimates into economic surplus models. See Appendix 3 for summary tables of all impact studies.

Figure 2. Evidence of better-identified causal effects

Study	Change	Effect
Bulte et al., 2014	Adoption of improved cowpea	0-9% increase in yield
Finkelstein et al., 2015	Consumption of iron-fortified pearl millet for four months	Statistically significant positive effects on three measures of iron-deficiency
Michler et al., 2018	Adoption of improved chickpea	23-28% increase in household profits
Verkaart et al., 2017	10% increase in area planted with improved chickpea	12.6% increase in household income; 3.9% reduction in probability of household poverty
Ndjeunga et al., 2013	Adoption of improved groundnut	155-202 kg/ha increase in yields
Simtowe et al., 2012	Adoption of improved groundnut	2% increase in household consumption expenditure; 12-17% decrease in poverty headcount

The available estimates of state- and country-level welfare impacts, and ROI are largely invalidated by their failure to identify farm-level causal effects of adoption. Interestingly, all of these studies find positive welfare impacts and ROI. While this may be a result of researchers' confirmation bias, it may also offer suggestive evidence (aligning with theory and intuition) that investments in crop improvement and agricultural development improve welfare and yield competitive returns. Improvements in future welfare and ROI analyses may be realized by incorporating better-identified farm-level impact estimates into economic surplus models. See Appendix 3 for summary tables of all impact studies.

4. Conclusions and recommendations

GLDC crops can contribute to the sustainable intensification of smallholder agriculture in sub-Saharan Africa and South Asia. In experimental settings, improved GLDC crops offer higher yields, increased drought and disease resistance, greater marketability, and better nutrition than traditional cereals such as maize, wheat, and rice. Intercropping and rotation of GLDC varieties with cereals may rebuild soil fertility through legume nitrogen fixation, reducing the need for chemical fertilizer inputs and generating positive spillovers into cereal yields. Nonetheless, the empirical evidence base on the impacts of improved GLDC crops in the field is, for the most part, poorly identified and incomplete. The objective of this review is therefore to synthesize the empirical evidence on improved GLDC crops in sub-Saharan Africa and South Asia, critically review the findings and methodologies of GLDC impact studies, and provide methodological recommendations for future analyses.

Based on keyword searches, we identified 18 studies measuring impacts from improved GLDC varieties in sub-Saharan Africa or South Asia since 2012 (see Appendix 2 for reviews of each study). These studies employ methodological approaches and datasets of varying quality. While all measure positive impacts from adoption or investment in improved GLDC crops, many fail to control for classic endogeneity problems such as simultaneity, omitted variable bias, and measurement error. Appendix 1 explores impact evaluation methodologies such as randomized experiments, quasi-experimental methods such as difference-in-difference and regression discontinuity design, and observational methods such as PSM, IV, and fixed effects. Appendix 1 also discusses good practices in economic surplus modeling and highlights the shortcomings of studies based on correlations and expert interviews.

The literature contains five reasonably well-identified causal impact evaluations of improved GLDC crop adoption (see Figure 2). These studies evaluate the effects of adoption on household yields, profits, income, consumption expenditure, and poverty; all estimate significant, positive treatment effects.[4] A further study measures the effects of iron-fortified pearl millet on schoolchildren's health, finding significant, positive treatment effects based on three measures of iron deficiency, illustrating potential nutritional benefits of GLDC crops. Macro-level welfare analyses and calculations of research ROI invariably estimate significant benefits from research and development (R&D) investment in improved GLDC crops. Nonetheless, most of these studies are invalidated by poorly identified farm-level effects. Further studies rely on correlations and expert interviews, leaving them at risk of selection bias and measurement error among other problems.

The majority of studies focus on chickpea and groundnut, to the detriment of other regionally important crops such as millet, cowpea, and soybean. Additionally, most studies concentrate on

---

[4] Bulte et al. (2014) estimate a treatment effect bounded between 0% and 9% in their double-blind procedure. In reference to poverty, a "positive" treatment effect is a reduction in poverty or in the probability of poverty.

Ethiopia, India, and Tanzania, with the majority of countries in the region left unevaluated. Given the high levels of heterogeneity in African agro-ecologies, more studies are needed to address plausible local variation in impacts. Finally, nearly all impact studies seek to measure economic impacts, neglecting environmental, nutritional, and social impacts. Future impact evaluations could address this imbalance.

4.1 Recommendations for future impact evaluations

The preceding synthesis highlights a number of areas for future improvements in the evaluation of GLDC crops. These improvements may derive from the incorporation of new technologies and methodologies, and may yield more accurate and nuanced estimates of GLDC impacts, potentially guiding policy.

4.1.1 Improving the scope and quality of data collection

Data collection through field surveys is subject to substantial measurement error due to enumerator error, recall error, and difficulty in accurately observing key variables. Measurement error can be mitigated by incorporating promising new technologies including: DNA fingerprinting to identify crop varieties, remote sensing (including drone and satellite imagery) to measure crop area and other variables, and tablets and cellphones for data entry (see Strengthening Impact Assessment in the CGIAR, 2017). Some studies reviewed above have already made use of these technologies, but the scope to expand their application is evident.

To reduce potential bias from omitted variables, it is important to control for as many confounding variables as possible, particularly in the context of non-experimental designs. When appropriate to the context, field surveys could gather data on soil and land quality in relation to improved varieties, as well as effort indicators such as labor inputs in order to control for adoption-induced input reallocation (Bulte et al., 2014). Another potentially important – yet often omitted – variable is farmers' aspirations, which shape their on-farm preferences and decision making. Mausch et al. (2018) and Verkaart et al. (2018) outline early approaches to measuring farmers' aspirations.

Many studies measure the impacts of adoption on yields within a single field. In practice however, households optimize total income or profit, not yields. Total income may derive from complex and diversified livelihood strategies, of which improved crop cultivation may constitute only a small part. Thus, it is theoretically more coherent to measure the impacts of improved crop adoption on total household income or profits, sidestepping potential bias resulting from intra-household input reallocation (Michler et al., 2018).

4.1.2 Building impact evaluation into programme implementation

In cases where programme implementation is "targetable" (e.g. through demonstration plots, training sessions and seed distribution), randomized implementation may not be significantly more costly than non-random assignment. In these cases, effective randomization may offer a valid IV for adoption and enable identification of adoption effects (see 2.3.2). Thus, pre-rollout planning should at least consider the potential for randomized implementation. In cases where

randomization is infeasible, non-experimental designs (e.g. difference-in-differences) can be strengthened when planned in advance.

Effective randomization may also remove the need to conduct a baseline survey, reducing data collection costs. In fact, researchers may be able to achieve more robust results by conducting multiple follow-up surveys of a smaller sample than by conducting baseline and follow-up surveys with a larger sample (McKenzie, 2012).

Recent work by Hakizimfura et al. (2018) highlights the possibility of oversampling compliers in contexts of low compliance with treatment. If there is concern that a treatment will have particularly low uptake, or that treatment will substantially increase variance in outcomes for the treatment group, oversampling can increase sample power more cost-effectively than increasing overall sample size. Likely compliers may be identified by estimating regressions over pre-existing or pre-experimental data, and focusing on highly correlated covariates. Final analyses can be reweighted by the probability of being sampled in order to recalibrate sample representativeness (McKenzie, 2018).

Another recent methodological innovation relevant to cost-constrained practitioners is proposed by Carneiro et al. (2017). These authors offer a statistical procedure for optimally selecting which covariates to include in baseline and follow-up surveys. Precision and power may be improved by either increasing sample size or collecting more relevant covariates, but budget constraints require tradeoffs between these two costly efforts. Carneiro et al.'s method therefore works by applying a "greedy" algorithm to determine whether an additional covariate improves power enough to justify additional data-collection costs. The method results in optimized surveys that can achieve increased sample power and average treatment effect precision at substantially lower data-collection costs (Holla, 2018; Carneiro et al., 2017).

It is also critical to bear in mind the important insights provided by Bulte et al. (2014) (reviewed above). Even within the context of a well-run RCT, isolating the effects of adoption from other proximal effects (e.g. management effort and uptake of improved management practices) induced by promotional efforts may not be straightforward. In other words, randomized assignment of promotion does not automatically meet the conditions of the exclusion restriction in the case of adoption. The promotional effort may bring about one or more intermediary effects (e.g. improved skills) that are correlated with both adoption and the outcomes of interest. It is therefore important to complement randomized adoption studies with efforts to interrogate the mechanisms linking treatment (the improved crop promotional intervention) and any evidence of outcomes (see for example Hughes et al., 2018.)

4.1.3 Evaluating non-economic impacts
The existing literature is heavily focused on economic outcomes, leaving environmental, nutritional, and social impacts largely unassessed. Future studies focused on nutrition and environment may need to borrow methodological approaches from the health and environmental assessment literatures. Assessments of the social impacts of holistic farm-level

interventions (which may include promotion of improved GLDC crops as a component) could include analyses of intra-household bargaining (see for example Doss, 2013), farmers' aspirations and livelihood strategies (see Mausch et al., 2018), and migration (Quinn, 2009; Mendola, 2008).

4.1.4 Rethinking the objective of impact studies

Impact studies that estimate positive results may ensure continued financial support from governments and donors. Researchers may therefore feel significant pressure to arrive at findings favorable to specific programmes or projects. At the same time, the utility of impact assessments arises from their reliable identification of what is working and not working in the field. Thus, it is of the utmost importance that researchers be encouraged and enabled to conduct impact assessments free from internal or external pressures and interests. The objective of impact assessment should be to establish an evidence base that informs policymaking – not to confirm or legitimize favoured programmes.

5.0 Referenced impact studies

Asfaw, Solomon, Bekele Shiferaw, Franklin Simtowe, and Leslie Lipper, Impact of Modern Agricultural Technologies on Smallholder Welfare: Evidence from Tanzania and Ethiopia, *Food Policy*, Vol. 37, pp. 283-295 (2012).

Bantilan C, Charyulu D. Kumara, P.M. Gaur, M.D. Shyam, and D. Jeff, *Short-Duration Chickpea Technology: Enabling Legumes Revolution in Andhra Pradesh, India*, Research Report No. 23. ICRISAT, 208 pp. (2014).

Bulte, Erwij, Gonne Beekman, Salvatore Di Falco, Joseph Hella, and Pan Lei, *Behavioral* Responses and the Impact of New Agricultural Technologies: Evidence from a Double-Blind Field Experiment in Tanzania, *American Journal of Agricultural Economics*, Vol. 96, No. 3, pp. 813-830 (2014).

Finkelstein, Julia, Saurabh Mehta, Shobha A Udipi, Padmini S Ghugre, Sarah V Luna, Michael J Wenger, Laura E Murray-Kolb, Eric M Przybyszewski, and Jere D Haas, A Randomized Trial of Iron-Biofortified Pearl Millet in School Children in India, *Journal of Nutrition* (2015).

Gierend, Albert, Henry Ojulong, and Nelson Wanyera, *A Combined Ex-Post/Ex-Ante Impact Analysis for Improved Sorghum and Finger Millet Varieties in Uganda*, ICRISAT, Socioeconomics Discussion Paper Series No. 19 (2014).

Kaliba, A, K. Mazvimavi, and G. Ghebreyesus. Economic Profitability and Risk Analyses of Improved Sorghum Varieties in Tanzania, *Journal of Development and Agricultural Economics*, Vol. 9, No. 9, pp. 250-268 (2017).

Karunakaran K.R., C.P. Gracy, H. Lokesha, C. Bantilan, Charyulu D. Kumara, Rao P Parthasarathy, Rao G.D. Nageswara, M. Vaithiyalingan, H.L. Nadaf, P. Venkataramana, H.D. Upadhyaya, P. Janila, and K.P.C. Rao, *Groundnut Baseline and Early-Adoption Surveys in South Asia Insights from TL-II (Phase-1) Project: Synthesis Report*, ICRISAT, 104 pp (2013).

Kumara, Charyulu D, Shyam D. Moses, Cynthia Bantilan, S.T. Borikar, S.K. Gupta and K.N. Rai, *Pearl Millet Technology Adoption and Impact Study in Maharashtra*, Research Report 71, ICRISAT, 76 pp. (2016).

Kumara Charyulu D, Shyam D Moses, Cynthia Bantilan, S.T. Borikar, A. Ashok Kumar and Belum V.S. Reddy, *Rainy Season Sorghum Technology Adoption and Impact Study in Maharashtra*, Research Report 70, ICRISAT (2016).

Kumara Charyulu D, Cynthia Bantilan, A. Rajalaxmi, B.V.S. Reddy, S.T. Borikar, Kumar A. Ashok, N.P. Singh and Shyam D. Moses, *Development and Diffusion of Sorghum Improved Cultivars in India: Impact on Growth and Stability in Yield*, Working Paper Series No. 50. 92 pp. (2014).

Macharia, Ibrahim, Alastair Orr, Franklin Simtowe, and Solomon Asfaw, *Potential Economic and Poverty Impact of Improved Chickpea Technologies in Ethiopia*, International Association of Agricultural Economists Triennial Conference (2012).

Michler, Jeffrey, Emilia Tjernström, Simone Verkaart, and Kai Mausch, Money Matters: The Role of Yields and Profits in Agricultural Technology Adoption, *American Journal of Agricultural Economics*, pg. 1-22 (2018).

Ndjeunga, J., B.R. Ntare, H. Ajeigbe, C.A. Echekwu, A. Ibro, and A. Amadou, *Adoption and Impacts of Modern Groundnut Varieties in Nigeria*, ICRISAT (2013).

Nwankwo Chianu, Jonas, Jeroen Huising, Seth Danso, Peter Okoth, Justina Nwanganga Chianu, and Nteranya Sanginga, Financial Value of Nitrogen Fixation in Soybean in Africa: Increasing Benefits for Smallholder Farmers, *Journal of Life Sciences*, Vol. 4, No. 31 (2010).

Shiferaw, Bekele, Tewodros Kebede, and Liang You, Technology Adoption under Seed Access Constraints and the Economic Impacts of Improved Pigeonpea Varieties in Tanzania, *Agricultural Economics*, Vol. 39, pp. 309-323 (2008).

Simtowe, Franklin, Menale Kassie, Solomon Asfaw, Bekele Shiferaw, Emmanuel Monyo, and Moses Siambi, *Welfare Effects of Agricultural Technology adoption: the Case of Improved Groundnut Varieties in Rural Malawi*, International Association of Agricultural Economists Triennial Conference (2012).

Tsusaka, Takuji, Harry Msere, Moses Siambi, Kizito Mazvimavi, and Patrick Okori, *Evolution and Impacts of Groundnut Research and Development in Malawi: An Ex-post Analysis*, Vol. 11(3), pp. 139-158 (2016).

Verkaart, Simone, Bernard Munyua, Kai Mausch, and Jeffrey Michler, Welfare Impacts of Improved Chickpea Adoption: A Pathway for Rural Development in Ethiopia? *Food Policy*, Vol. 66, pp. 50–61 (2017).

Additional references

Alston, Julian, George Norton, and Philip Pardey, Science Under Scarcity: Principles and Practice for Agricultural Research Evaluation and Priority Setting, *American Journal of Agricultural Economics*, Vol. 77, No. 4 (1995).

Angrist, Joshua and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, 1st Edition, Princeton University Press (2009).

Angrist, Joshua and Alan Krueger, Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments, *Journal of Economic Perspectives*, Vol. 15, No. 4, pg. 69-85 (2001).

Barrett, Christopher and Michael Carter, The Power and Pitfalls of Experiments in Development Economics: Some Non-Random Reflections, *Applied Economic Perspectives and Policy*, Vol. 32, No. 4, pp. 515-548 (2010).

Bruhn, Miriam and David McKenzie, In Pursuit of Balance: Randomization in Practice in Development Field Experiments, *American Economic Journal: Applied Economics*, Vol. 1, pp. 200-232 (2009).

Bulmer, Martin and Donald Warwick, Social Research in Developing Countries: Surveys and Censuses in the Third World, *Journal of the Royal Statistical Society*, Vol. 157, No. 2 (1994).

Caliendo, Marco and Sabine Kopeinig, *Some Practical Guidance for the Implementation of Propensity Score Matching*, IZA Discussion Paper No. 1588 (2005).

Carneiro, Pedro, Sokbae Lee, and Daniel Wilhelm, *Optimal Data Collection for Randomized Controlled Trials*, Working Paper, University College London (2017).

CGIAR Research Programme, *Grain Legumes and Dryland Cereals Agri-Food Systems*, Proposal (2018).

Clarke, Damian, *Estimating Difference-in-Differences in the Presence of Spillovers*, Munich Personal RePEc Archive, Paper No. 81604 (2017).

Cook, Thomas, William Shadish, and Vivian Wong, Three Conditions Under which Experiments and Observational Studies Produce Comparable Causal Estimates: New Findings From Within-study Comparisons, *Journal of Policy Analysis and Management* (2008).

Deaton, Angus and Nancy Cartwright, Understanding and Misunderstanding Randomized Controlled Trials, *Social Science and Medicine*, Vol. 210, pp. 2-21 (2018).

Deaton, Angus, *Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development*, NBER Working Paper No. 14690 (2011).

De Janvry, A., G. Graff, E. Sadoulet, and D. Zilberman, *Technological Change in Agriculture and Poverty Reduction*, Concept paper for the World Development Report on Poverty and Development (2001).

Dercon, Stefan and Douglas Gollin, Agriculture in African Development, Theories and Strategies, *Annual Review of Resource Economics*, Vol. 6, pg. 471-492 (2014).

Djimeu, Eric and Deo-Gracias Houndolo, *Power Calculation for Causal Inference in Social Science*, Working Paper 26, International Initiative for Impact Evaluation (2016).

Doss, Cheryl, *Intrahousehold Bargaining and Resource Allocation in Developing Countries*, World Bank Policy Research Working Paper (2013).

Duflo, Esther, Rachel Glennerster, and Michael Kremer, *Using Randomization in Development Economics Research: A Toolkit*, Center for Economic Policy Research, Discussion Paper Series No. 6059 (2007).

Gertler, Paul, Sebastian Martinez, Patrick Premand, Laura Rawlings, and Christel Vermeersch, *Impact Evaluation in Practice*, 2nd Edition, World Bank Publications (2016).

Ghosh, P. K., K. K. Bandyopadhyay, R. H. Wanjari, M. C. Manna, A. K. Misra, M. Mohanty, and A. Subba Rao, Legume Effect for Enhancing Productivity and Nutrient Use-efficiency in Major Cropping Systems – An Indian Perspective: A Review, *Journal of Sustainable Agriculture*, Vol. 30, No. 1, pg. 59-86 (2007).

Gumma, Murali Krishna, Irshad Mohammed, and Anthony Whitbread, *Characterization of GLDC Mega-environments*, ICRISAT (2017).

Hakizimfura, Emmanuel, Douglas Randall, and Bilal Zia, *Decentralized Delivery of Financial Education*, Policy Research Working Paper No. 8521, World Bank (2018).

Holla, Alaka, *Are We Overinvesting in Baselines?* Development Impact Blog, World Bank (2018).

Hughes K, S. Morgan, K. Baylis, J. Oduol, E. Smith-Dumont, T.G. Vågen, and H. Kegode, *Assessing the Downstream Socioeconomic Impacts of Agroforestry in Kenya.* ICRAF Working Paper No. 291. World Agroforestry (2018). DOI: http://dx.doi.org/10.5716/WP18033.pdf.

Huho, Julius and Edward Mugalavai, The Effects of Drought on Food Security in Kenya, *International Journal of Climate Change Impacts and Responses*, Vol. 2, No. 2 (2010).

Imbens, Guido, Matching Methods in Practice: Three Examples, *Journal of Human Resources*, Vol. 50, pp. 373-419 (2015).

Imbens, Guido and Thomas Lemieux, Regression Discontinuity Designs: A Guide to Practice, *Journal of Econometrics*, Vol. 142, pp. 615-635 (2008).

International Food Policy Research Institute, *Dynamic Research Evaluation for Management (DREAM) User's Manual*, Version 3 (2001).

Jayne, Thomas, Jordan Chamberlin, and Derek Headey, Land Pressures, the Evolution of Farming systems, and Development Strategies in Africa, *Food Policy*, Vol. 48, pp. 1-17 (2014).

King, Gary and Richard Nielsen, *Why Propensity Scores Should not be Used for Matching*, Working Paper, Institute for Quantitative Social Science, Harvard University (2016).

Lousdal, Mette Lise, An Introduction to Instrumental Variable Assumptions, Validation, and Estimation, *Emerging Themes in Epidemiology*, Vol. 15 (2018).

Mausch, K., D. Harris, E. Heather, E. Jones, J. Yim, and M. Hauser, Households' Aspirations for Rural Development Through Agriculture, *Outlook on Agriculture*, pp. 1-8 (2018).

Mayes, S., F.J. Massawe, P.G. Alderson, J.A. Roberts, S.N. Azam-Ali, and M. Hermann, The Potential for Underutilized Crops to Improve Security of Food Production, *Journal of Experimental Botany*, Vol. 63, No. 3, pp. 1075-1079 (2012).

McKenzie, David, *Should you Oversample Compliers if Budget is Limited and you are Concerned Take-up is Low?* Development Impact Blog, World Bank (2018).

McKenzie, David, Beyond Baseline and Follow-up: The Case for More T in Experiments, *Journal of Development Economics*, Vol. 99, pp. 210-221 (2012).

Mendola, Mariapia, Migration and Technological Change in Rural Households: Complements or Substitutes*? Journal of Development Economics*, Vol. 85, pp. 150-175 (2008).

Pardey, Philip, Julian Alston, and Roley Piggott*, Agricultural R&D in the Developing World*: *Too Little, Too Late?* International Food Policy Research Institute (2006).

Quinn, Michael, Estimating the Impact of Migration and Remittances on Agricultural Technology, *Journal of Developing Areas*, Vol. 43, No. 1, pp. 199-216 (2009).

Roe, Brian and David Just, Internal and External Validity in Economics Research: Tradeoffs between Experiments, Field Experiments, Natural Experiments and Field Data, *American Journal of Agricultural Economics*, Proceedings (2009).

Rosenbaum, Paul and Donald Rubin, The Central Role of the Propensity Score in Observational Studies for Causal Effects*, Biometrika*, Vol. 70, No. 1, pp. 41-55 (1983).

Sarsons, Heather, Rainfall and Conflict: A Cautionary Tale, *Journal of Development Economics*, Vol. 115, pp. 62-72 (2015).

Strengthening Impact Assessment in the CGIAR, *Final Report*, Independent Science and Partnership Council (2017).

Sunding, David and David Zilberman, The Agricultural Innovation Process: Research and Technology Adoption in a Changing Agricultural Sector, *Handbook of Agricultural Economics*, Vol. 1, Part A, pp. 207-261 (2001).

Valbuena, Diego, Olaf Erenstein, Sabine Homann-Kee Tui, Tahirou Abdoulaye, Lieven Claessens, Alan J. Duncan, and Bruno Gérard, Conservation Agriculture in Mixed Crop-livestock Systems: Scoping Crop Residue Trade-offs in Sub-Saharan Africa and South Asia, *Field Crops Research*, Vol. 132, pp. 175-184 (2012).

Verkaart, Simone, Kai Mausch, and Dave Harris, Who are Those People we Call Farmers? Rural Kenyan Aspirations and Realities, *Development in Practice*, Vol. 28, No. 4, pp. 468-479 (2018).

Wooldridge, Jeffrey, *Econometric Analysis of Cross-Section and Panel Data*, 2nd Edition, MIT Press (2010).

World Bank, *World Development Report 2008: Agriculture for Development* (2008).

Appendix 1. Approaches to impact evaluation in agriculture

It is important that policymakers understand the degree to which adoption of improved GDLC crops improves farmers' yields, profits, incomes, nutrition, food security, and soil fertility. Experts at the institution and country levels need to know whether investments in R&D of improved GLDC crops are cost effective relative to alternative investments. Obtaining accurate answers to these questions presents a significant challenge, but one for which the field of impact evaluation offers a range of methodological approaches. This section outlines these approaches and provides references to more complete treatments.

Impact evaluation is fundamentally concerned with causal inference, whereby the researcher seeks to measure outcomes for "treated" individuals[5] relative to what the outcomes would have been for these individuals if the treatment had never occurred. Since it is impossible to observe the alternate universe in which treated individuals went untreated (known as the Fundamental Problem of Causal Inference), the "missing counterfactual" must be proxied by experimental construction, or assumptions regarding the data-generating process, in order to identify true causal treatment effects (Angrist and Pischke, 2009).

Impact evaluation is beset by problems of endogeneity, which arise when any explanatory variables in a regression model are correlated with the error term. Endogeneity is an issue when treatment is not "as-if-randomly assigned" (Wooldridge, 2010). Endogeneity typically arises from simultaneity, omitted variables, and measurement error.

Methodological approaches to addressing these endogeneity problems include randomized experiments, quasi-experimental methods such as difference-in-differences and regression discontinuity design, and observational methods such as propensity score matching (PSM), instrumental variables (IVs), and fixed effects (Gertler, 2016). Other methods such as correlational studies, expert interviews, and opinion surveys, can inform researchers' knowledge of the data-generating process and indicate areas for further inquiry, but are incapable of identifying causal impacts.

1.1 Experimental approach

The simplest impact-evaluation design that allows for valid causal inference is the randomized experiment or randomized control trial (RCT). Random assignment of treatment ensures that treatment is uncorrelated with individuals' characteristics, effectively creating a counterfactual control group that proxies the "alternative universe" where the treated went untreated. Randomization largely mitigates the endogeneity problems identified above, and in its simplest form allows calculation of average treatment effects as simple comparisons of sample means for treatment and control groups (Angrist and Pischke, 2009).

---

[5] "Individuals" in this case could be farmers, households, villages, firms, etc. The generic term "treatment" could refer either to an experimentally assigned treatment such as a randomized grant of improved seeds, or the endogenously determined adoption of improved seeds.

RCTs have received substantial attention in the agricultural development literature over recent decades. They are often treated as the "gold standard" of impact evaluation methodologies. While this is true in many contexts, RCTs do have important limitations (Deaton, 2009; Deaton and Cartwright, 2018). First, RCTs cannot be conducted in retrospect: they require prior planning and careful implementation, and are therefore unusable for the evaluation of any policy that has already been implemented in a non-randomized fashion. Many important policies (especially larger-scale policies) are implemented universally or non-randomly, which highlights the importance of possessing alternative quasi-experimental and observational impact evaluation methodologies.

Second, RCTs require that treatment be administered in a precise, controlled fashion. This may be feasible in the case of demonstration plots, training interventions, or improved seed distribution, but may be more difficult for across-the-board variety releases and non-targeted promotion. If agricultural research institutions intend to build more randomization into their policy implementation, they may first need to strengthen the targeting and precision of that implementation.

Third, precise estimation of average treatment effects in an experimental setting requires blinding or double-blinding to avoid placebo effects. Nonetheless, there are ethical limitations to offering placebo treatments in social contexts (one cannot, for instance, offer fake microloans or mosquito nets). See further discussion of ethical dilemmas regarding RCTs in Barrett and Carter (2010).

Thus, while RCTs are an essential tool for impact evaluation, they should be considered as one part of a larger toolkit. In some cases, the advantages of observational studies (i.e. larger sample sizes, avoidance of blinding-problems and non-compliance) may outweigh the potential identification problems inherent in non-experimental methods (Deaton and Cartwright, 2018). Furthermore, under certain conditions, observational studies appear to arrive at quantitatively comparable estimates to those found in randomized experiments (Cook et al., 2008). Nonetheless, these cautionary notes should not suggest that experiments not be undertaken when feasible. Randomization may be built into implementation for little additional cost, and effective randomization may even obviate the need for a baseline survey (McKenzie, 2012).

1.2 Quasi-experimental approaches

At times, individualized random assignment of treatment is not feasible, as is the case when a large-scale policy is rolled out for an entire population. This may occur, for instance, when an agricultural research institution develops and releases an improved GLDC variety within the boundaries of one state, but not neighbouring states. Despite the lack of random assignment of treatment, researchers may still desire to estimate the causal effect of the improved variety on, say, farmers' yields and profits. One set of approaches to this causal inference problem is often

referred to as quasi-experimental. Quasi-experimental approaches are distinguished by careful ex-ante research design and data collection focused on estimating a causal effect within a specific time and setting (Cook et al., 2008).

1.2.1 Difference-in-difference estimation

When treatment is non-randomly assigned, it is likely that there exist both observable and unobservable differences between treated and untreated populations. For instance, farmers who adopt an improved variety are likely to be systematically different than non-adopters across both observables (e.g. farm size, education, location) and unobservables (e.g. ability, motivation). In this case, simply comparing outcomes for "treated" and "untreated" individuals (adopters and non-adopters) will offer a biased estimate of the effects of adoption. Difference-in-difference estimation, in contrast, enables the construction of a valid counterfactual by imposing the assumption that, absent treatment, both observed and unobserved differences between adopters and non-adopters would be constant (parallel) over time. In other words, the assumption is that a time-invariant fixed effect is all that differs between the two groups (Wooldridge, 2010).

This may often be a reasonable assumption, since farm size, education, location, ability, and motivation do not often change substantially over time. Given this "parallel trends assumption", a treatment effect may be estimated from the difference between pre- and post-treatment differences between treated and un-treated groups – that is, a difference-in-differences.[6] This method requires data on both treated and untreated individuals at both the pre- and post-treatment stages, which means that panel data collection is required. Preferably, more periods would be included in the panel in order to visually verify the parallel trends assumption (Angrist and Pischke, 2009).

1.2.2 Regression discontinuity design

Difference-in-difference estimation departs from the assumption that treatment and control groups are different across both observables and unobservables. In contrast, regression discontinuity design (RDD) departs from the identifying assumption that some boundaries are cutoffs for determining treatment but are in all other senses arbitrary, making individuals on either side of the boundary comparable across all covariates except the treatment ("as-if-random" treatment). RDD was initially applied largely for cutoffs such as test scores and age brackets, but has also been used to assess treatment effects across arbitrary geographic boundaries (Imbens and Lemieux, 2008). This method may therefore be relevant in cases where a "treatment" such as certification, release of an improved variety, or a state-wide promotion programme is implemented in one geographic region, but not in neighbouring geographic regions that are otherwise similar. In such a situation, treated individuals near a border may differ from their counterparts on the other side of the border only in terms of treatment. In practice, data would be collected from both sides of the border before and after treatment in order to verify

---

[6] Technically, the identifying assumption underlying difference-in-differences is the Stable Unit Treatment Value Assumption (SUTVA), which requires that no spillovers occur between quasi-treatment and quasi-control groups. Clarke (2017) proposes a method for dealing with local violations of the SUTVA.

the pre-treatment balance. Distance from the border would be included as a control (see Angrist and Pischke [2009] for a more complete explanation).

There are numerous challenges arising from RDD in this context. First, most geographical borders are not arbitrary, but demarcate meaningful institutional, ethnic, or economic differences. This suggests that arbitrary borders (i.e. between agricultural research jurisdictions) would be more effective than state boundaries. Practitioners could even roll out policy implementation over arbitrarily defined regions in order to construct their own RDD. Second, the RDD identifying assumption holds most easily when the division is "sharp," meaning that cross-border contamination is negligible. In the case of improved variety release or promotion, there is likely to be substantial cross-border seed trading and selling, making it a case of "fuzzy" RDD, which requires additional methods of control (IVs). Finally, since treatment effects estimated through RDD are generated under such locally specific conditions, they suffer from limited external validity (Imbens and Lemieux, 2008). Nonetheless, thinking about impact evaluation through the RDD lens may reveal useful natural experiments that researchers may otherwise have overlooked.

1.3 Other methods of causal inference

Further approaches to causal inference – best referred to as observational methods – do not require prior data collection or extensive ex-ante research design, but nonetheless may allow valid identification under certain conditions. Thus, these methods are often the most widely applicable in field settings.

1.3.1 Propensity score matching

In non-experimental settings, it is likely that farmers choosing to adopt an improved variety are systematically different from farmers who do not adopt it. These systematic differences lead to selection bias, whereby the naively estimated impact of adoption on outcomes is likely inflated by the effects of unobserved covariates that influence adoption. In other words, the simple counterfactual of non-adopters is not an accurate comparison for adopters. As proposed in Rosenbaum and Rubin (1983), it may be possible to construct a valid counterfactual by matching treated individuals (adopters) with non-treated individuals (non-adopters) who are as similar as possible to each other across all covariates besides the treatment. Estimation of propensity scores, which measure the probability that an individual with certain characteristics receives treatment, reduces a vector of observable covariates to a scalar score, which may then be used to match treated and non-treated individuals (Imbens, 2015).

1.3.2 Instrumental variables

Instrumental variable (IV) approaches are some of the most widely used methods of controlling for endogeneity resulting from measurement error, omitted variables, and simultaneity. See Angrist and Krueger (2001) for a deeper explanation of their development and applications.

1.3.3 Panel data (fixed effects)

Unobserved time-invariant heterogeneity may be controlled for in panel data with the use of fixed effects, which differentiate variables that are constant over time. Since we may suspect that important contributors to selection bias, such as ability, are relatively constant over time, fixed effects may be able to substantially reduce endogeneity.

1.4 Macro-level welfare analyses and returns on research investment

Many studies within the GLDC impact evaluation literature attempt to extrapolate welfare impacts from GLDC crop adoption to the state or national level. Many of these studies also seek to estimate economic returns from – and cost-effectiveness of – agricultural R&D programmes. Studies of this nature proceed by estimating individual-level yield, income, or profit gains from adoption relative to a counterfactual, and then extrapolating this gain by the extent of adoption in the sample area – a method outlined in Alston et al. (1995). In practice, most welfare and return-on-research-investment (ROI) analyses of this kind generate estimates through use of single-commodity, partial equilibrium, multi-market economic surplus models such as the Dynamic Research Evaluation for Management (DREAM) model developed by the International Food Policy Research Institute (IFPRI). These models enable researchers to input key data on yield gains, adoption, R&D costs, and supply and demand elasticities in order to arrive at estimates of economic surplus generated by R&D investments (IFPRI, 2001).

As with all modeling methods however, the quality of data input determines the quality of data output. Specifically, the accuracy of the total welfare and ROI estimates depends entirely on: (i) well-identified estimates of the individual-level yield/income/profit impact of the technology; (ii) the applicability of these estimates to broader contexts (external validity); (iii) the quality of data on R&D costs; and 4) the validity of claims regarding alternative investment options.

1.5 Correlational studies, expert interviews, and opinion surveys

Studies based on correlations between adoption and yields, incomes, profits, etc. are afflicted by the endogeneity problems highlighted above (including simultaneity, omitted variables, and measurement error). These problems likely lead to substantial bias in the estimated impacts of adoption. Correlations should not be considered evidence of impacts.

Ndjeunga et al. (2013) posit that expert interviews may constitute a low-cost alternative to field surveys for the assessment of adoption. Indeed, numerous studies in the GLDC adoption and impact evaluation literature use expert interviews to estimate extent of adoption and R&D costs. Nonetheless, when Ndjeunga et al. compare estimates of adoption based on expert interviews with estimates based on household surveys, they find wide differences, suggesting that expert interviews are *not* a reliable means of estimating these variables. The poor quality of expert

interviews is likely a result of geographic heterogeneity, experts' non-representative interactions with farmers, and biased heuristics.

Finally, studies often cite opinion surveys of farmers as evidence that a specific policy is popular or successful. These surveys are subject to substantial bias from the following sources (Bulmer and Warwick, 1993):

- o *Enumerator influence*: respondents may feel pressured to report positively on a programme with which the enumerator is associated.
- o *Recall error/bias*: Respondents may recall past data in a biased manner due to recall difficulty and present bias.
- o *Subjective responses*: Many survey questions are subjective in nature, resulting in arbitrary variation in responses due to heterogeneity in respondents.

Of course, some of these challenges are associated with quantitative field surveys as well, and the same good practices regarding sample selection and enumeration should be followed (See Bulmer and Warwick, 1993).

Appendix 2: Summaries and critiques of GLDC impact studies

This appendix compiles and reviews the recent literature on impact studies of improved GLDC varieties in sub-Saharan Africa and South Asia. Studies included in the review were obtained from the archives of the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), and searches on Google Scholar and AgEcon Search for different combinations of the terms "impact assessment," "impact evaluation," "adoption," "scaling," "improved varieties," and countries, regions, and crops included in the Consultative Group on International Agricultural Research (CGIAR) GLDC programme mandate.[7] Given the rapidly changing nature of rural economies, environmental conditions, and improved varieties, only studies from 2012 or later were considered (with two exceptions).[8] This yielded 18 relevant studies, all of which are reviewed below by methodological approach. See Appendix 3 for summary tables of all impact studies reviewed below.

2.1 Experimental approaches

Bulte, Erwij, Gonne Beekman, Salvatore Di Falco, Joseph Hella and Pan Lei, Behavioral Responses and the Impact of New Agricultural Technologies: Evidence from a Double-Blind Field Experiment in Tanzania, *American Journal of Agricultural Economics*, **Vol. 96, No. 3, pp. 813-830 (2014).**

Summary: RCTs in the social sciences are rarely blinded or double-blinded, meaning that experimental subjects know whether they are in the treatment or control group. This knowledge may induce behavioural responses that confound simple computation of an average treatment effect (ATE). To measure the extent of these confounding factors, Bulte et al. (2014) evaluate the impact of improved cowpea varieties on harvests for a sample of farmers in Tanzania using both a traditional RCT design and a double-blind RCT to control for placebo effects. Farmers in the double-blind experiment are given indistinguishable seeds and told that they may or may not be of an improved variety. Data is drawn from a random sample of 583 households in Morogoro, Tanzania, with multiple survey waves conducted in 2011. The sample suffers from considerable attrition, although the authors demonstrate that this is balanced over observables.

The authors first compute a "naïve" ATE from the standard RCT, and then compute a corrected ATE from the double-blind experiment. The standard ATE indicates that improved cowpea seeds increase harvest volumes by 27% relative to local varieties; the double-blind RCT suggests that two thirds of this measured impact is in fact a placebo effect caused by treatment-induced effort.

---

[7] The GLDC programme's priority crops and countries were determined by the metrics of "poverty prevalence, agro-ecological alignment, value of crop production, foresight projections of significant demand and/or deficit in supply, *ex-ante* return on research investment, consideration of quality, market and environmental traits and alignment with stakeholder priorities". Priority crops include five grain legumes: chickpea, cowpea, pigeonpea, groundnut, and soybean; and three dryland cereals: sorghum, pearl millet, and finger millet. Priority countries include Burkina Faso, Ethiopia, India, Malawi, Mali, Mozambique, Myanmar, Niger, Nigeria, Sudan, Tanzania, Uganda, and Zambia (CGIAR, 2018).

[8] Shiferaw et al. (2008) is included to provide more coverage of pigeonpea and highlight a better-identified approach to economic surplus modeling. Nwankwo et al. (2010) is included to provide coverage of soybeans and environmental impacts, which do not appear elsewhere in the sample.

Observables such as land quality and labour input are able to explain a portion of this placebo effect, but the majority is not captured in the data and is ascribed to "unobservable effort induced by treatment". The authors conclude by highlighting the importance of controlling for observable effort variables in standard RCT designs and implementing double-blind designs when possible.

Critique: The findings in this study illustrate the potential misestimation resulting from placebo effects. These findings call into question the true nature of impacts from improved GLDC varieties, and more broadly the reliability of unblinded experimental designs in agriculture. If much of the measured impact of improved varieties derives from induced effort responses rather than inherent trait superiority, perhaps development initiatives should reconsider their prioritization of crop breeding *vis a vis* extension.[9]

Importantly, the study suffers from a range of potential problems. First, accurate measurement of cowpea harvests is complicated because significant value from cowpea is derived from continuous harvesting of leaves, whereas Bulte et al. only measure final grain harvests. Second, some effort induced by treatment, such as fertilizer application, may indeed constitute an optimal response to belief in an improved variety's potential, and could be considered an intrinsic component of the variety's impact. Third, the study's external validity extends only to one district in Tanzania, which may be an extreme case as the authors themselves acknowledge. Fourth, yields in a single field are likely not the best measurement of impact since households optimize over total income rather than field-specific yields (Michler et al., 2018). Finally, distribution of false improved seeds is ethically dubious at best, given that farmers invest real resources (time, fertilizer, quality land) based on their belief in these seeds. The applicability of double-blind designs in agricultural development is limited by these ethical constraints.

Finkelstein, Julia, Saurabh Mehta, Shobha A Udipi, Padmini S Ghugre, Sarah V. Luna, Michael J. Wenger, Laura E. Murray-Kolb, Eric M. Przybyszewski, and Jere D. Haas, A Randomized Trial of Iron-Biofortified Pearl Millet in School Children in India, *Journal of Nutrition* (2015).

Summary: This study implements a double-blind RCT in one Indian secondary boarding school in order to evaluate the health impacts of daily administered iron-fortified pearl millet (with unimproved pearl millet as a control) on schoolchildren. The study observes 246 schoolchildren over six months, with treatment and control switched at the four-month mark in response to clear positive health impacts from the treatment. The sampled school was specifically chosen for the study due to students' high rates of pre-trial iron-deficiency.

---

[9] Clearly, no drastic conclusions should be drawn based one study with limited external validity. There are theoretical and empirical reasons to believe that improved varieties have positive impacts on farmers' outcomes (see remaining reviews). After all, even Bulte et al.'s double-blind experiment measures a positive ATE from improved cowpea adoption bounded between 0% and 9%.

The authors measure statistically significant, positive treatment effects on three out of four measures of iron deficiency at the four-month mark, demonstrating that iron-fortified pearl millet contributes positively to students' nutritional outcomes, with especially pronounced effects measured for the most iron-deficient students.

Critique: The study's primary limitation is one of external validity, given its small sample size of just one school chosen non-randomly based on its students' low iron levels. Thus, estimated treatment effects may be interpreted as upper bounds of potential impacts from iron-fortified pearl millet on Indian secondary schoolchildren.

2.2 Other methods of causal inference

Michler, Jeffrey, Emilia Tjernström, Simone Verkaart, and Kai Mausch, Money Matters: The Role of Yields and Profits in Agricultural Technology Adoption, *American Journal of Agricultural Economics*, **pp. 1-22 (2018).**

Summary: Michler et al. draw on panel data (600 households surveyed in three waves between 2006-2014) to estimate the effects of improved chickpea adoption on yields and profits for farmers in Ethiopia. The authors begin by noting that chickpea adoption in the sample area is high, despite evidence that adoption does not substantially increase yields. In light of this observation, they theorize that it may be more consistent to focus on households' economic returns such as profits, rather than yields, since in practice farm households optimize over total household income rather than individual crops.

The authors begin by estimating yields as a function of adoption and a rich set of control variables using ordinary least squares (OLS), fixed effects models, and correlated random coefficient models (a generalized case of fixed effects models that allows unobservable household heterogeneity to vary with adoption). Observable household-level shocks are controlled by geo-coded rainfall data (for weather shocks) and data on input usage (for internal household shocks such as death or illness). Importantly, effects are measured at the household rather than field level, thus capturing potential input reallocation effects that could bias impact estimates (as in Bulte et al., 2014). Results demonstrate that when fixed effects and input controls are included, improved chickpea adoption has no significant effect on yields. The authors then estimate household profits as a function of adoption and controls. In this case, fixed effects and correlated random coefficient models reveal that improved chickpea adoption has a significant, positive effect on profits of between 23% and 28%.

Critique: This study illustrates effective panel data techniques, which enable the authors to control for unobservable time-invariant heterogeneity across households, reducing selection bias. Moreover, the authors effectively control for observable time-variant heterogeneity with rainfall data and input-usage data. However, they are unable to control for unobservable time-variant household shocks – a limitation of this approach.

The findings from this study illustrate the importance of measuring economic impacts such as income and profits rather than yields alone. In this case, an impact assessment measuring only yield impacts would have concluded that improved chickpea offers no benefits to farmers, while in reality it increases profits by between 23% and 28%. Evidently, the appropriate unit of impact analysis for smallholder agriculture is total household income or profits, since this is what households optimize. Finally, this study illustrates the importance of accounting for inputs such as labour, land quality, and fertilizer, since households reallocate these resources across crops in response to seed quality –potentially biasing impact estimates.

Verkaart, Simone, Bernard Munyua, Kai Mausch, and Jeffrey Michler, Welfare impacts of improved chickpea adoption: A pathway for rural development in Ethiopia? *Food Policy*, **Vol. 66, pp. 50-61 (2017).**

Summary: This study draws on observational panel data (three waves between 2006 and 2014) to estimate the effects of improved chickpea adoption on income and poverty in Ethiopia. The authors address endogeneity through a two-stage estimation procedure. In the first stage, they estimate the scale of chickpea adoption at the household level as a function of access to technology and improved seed, observable household characteristics, and village- and time-fixed effects. Since technology and seed access are endogenous, both are instrumented by distance measures (i.e. distance to market), which predict access but may only be associated with the scale of adoption through their effects on access. In the second stage, income and poverty are estimated as functions of the scale of chickpea adoption and other controls. Endogeneity within this model is controlled by instrumenting observed scale of chickpea adoption with the predicted scale of adoption from the first stage, which is strongly correlated with observed scale of adoption but not with individual household heterogeneity.

Results from this estimation procedure indicate that a 10% increase in area planted with improved chickpea is associated with a 12.6% increase in income per capita, and a 3.9% reduction in probability of falling below the US$2/day poverty line. The authors conclude that improved chickpea varieties represent a viable means of reducing poverty and increasing welfare in Ethiopia.

Critique: The identification assumptions underlying this estimation approach are that distance measures are valid IVs for seed and technology access, and that the predicted scale of adoption is a valid IV for observed scale of adoption. The latter seems plausible given that it misses individual household heterogeneity by construction. On the other hand, it seems likely that distance from market or road may be associated with important unobservable variables that may likewise influence adoption, calling into question the exclusion restriction. Despite these limitations, Verkaart et al. (2014) present reasonably reliable observational evidence that improved chickpea adoption improves welfare, especially when considered in light of corroborating studies.

Ndjeunga, J., B.R. Ntare, H. Ajeigbe, C.A. Echekwu, A. Ibro, and A. Amadou, *Adoption and Impacts of Modern Groundnut Varieties in Nigeria*, **ICRISAT (2013).**

Summary: Based on a random cross-sectional survey of 2,732 farming households from 245 randomly selected villages in semi-arid groundnut-producing regions of Nigeria in 2012, Ndjeunga et al. estimate the impacts of improved groundnut adoption on household yields, consumption, and revenues. Selection bias in crop adoption is addressed by means of PSM. Two alternative matching algorithms (nearest neighbour matching and kernel matching) are employed in order to pair each adopting household with a non-adopting household that is similar along observable covariates. Both matching algorithms generate reasonable balance over observables.

Two further efforts are made to control for potential selection bias over unobservables. First, observable covariates are lagged (based on recall) in order to reduce the influence of potential simultaneity between outcome variables and adoption. Second, an IV approach is incorporated to address potential selection bias from unobservables, with the predicted probability of awareness of improved varieties (itself a function of household and village characteristics) used to instrument adoption. The authors demonstrate that predicted awareness is a relevant IV for adoption, arguing that predicted awareness only affects outcome variables through its association with adoption.

Based on this identification strategy, the study estimates that adoption of improved groundnut varieties causes a 155-202 kg/ha increase in yields and a statistically significant increase in an index of food consumption, but has no measurable effects on revenues or poverty. Estimated effects are larger for wealthier farmers.

Critique: The study draws on a large, representative sample of farm households, and convincingly corrects for selection bias over observable covariates using PSM. Selection bias from unobservables may be mitigated by use of an IV approach, although it is unlikely that the instrument meets the exclusion restriction given that the covariates used to predict it are likely correlated with both outcomes and adoption. While imperfect, results based on this collection of estimation methods likely constitute an improvement on an endogenous OLS model.

Asfaw, Solomon, Bekele Shiferaw, Franklin Simtowe, Leslie Lipper, Impact of modern Agricultural Technologies on Smallholder Welfare: Evidence from Tanzania and Ethiopia, *Food Policy*, **Vol. 37, pp. 283-295 (2012).**

Summary: This study estimates the welfare impacts (measured in adult-equivalent units of consumption expenditure) of improved variety adoption in Ethiopia (chickpea) and Tanzania (pigeonpea). Observational data are drawn from cross-sectional surveys of 700 randomly selected households in chickpea-producing districts of Ethiopia, and 613 randomly selected households in pigeonpea-producing districts of Tanzania in 2008. The authors attempt to

overcome selection bias in the adoption variable by means of an endogenous switching regression, in which welfare functions are estimated separately and simultaneously for adopters and non-adopters based on the simultaneous estimation of a latent selection function using an IV approach.

Asfaw et al. argue that contact with government extension services and participation in farmer participatory variety selection programs (FPVS) are valid instruments in Ethiopia, and that contact with government and NGO extension services, and participation in FPVS programmes are valid instruments in Tanzania. They demonstrate that these variables do indeed predict adoption, satisfying the relevance assumption underlying IV. Yet they do not fully justify the exclusion restriction that these variables only affect consumption expenditure through their effect on adoption.

Based on this estimation strategy, results indicate that adoption of improved chickpea varieties is associated with a 24.6% increase in welfare for adopters and a hypothetical 20.9% increase in welfare for non-adopters in Ethiopia, while adoption of improved pigeonpea varieties is associated with a 103% increase in welfare for adopters and a hypothetical 99.4% increase in welfare for non-adopters.

Critique: The magnitude of estimated impacts raises substantial doubts regarding the reliability of these findings, especially in the case of pigeonpea in Tanzania. Furthermore, it seems implausible that the exclusion restriction holds in either case, since it is highly likely that participation in FPVS programmes and contact with extension agents are associated with household consumption levels. Extension agents more often visit wealthy, better-connected, or better-located farmers than they do poor, marginalized, or remote farmers. This means that the proposed IVs may be associated with the outcome variable through other mechanisms than their effect on adoption. Failure of the exclusion restriction results in an invalid IV estimation. The findings presented in this study should therefore be treated with caution.

Simtowe, Franklin, Menale Kassie, Solomon Asfaw, Bekele Shiferaw, Emmanuel Monyo, and Moses Siambi, *Welfare Effects of Agricultural Technology Adoption: The Case of Improved Groundnut Varieties in rural Malawi*, International Association of Agricultural Economists Triennial Conference (2012).

Summary: This study estimates the impact of adoption of improved groundnut varieties on household consumption expenditures and poverty indices in Malawi. Estimates are based on data drawn from a cross-sectional household survey conducted in 2008, which randomly sampled 594 households from 48 villages in groundnut-producing districts. PSM is employed to control for the endogeneity of adoption. The authors demonstrate that both nearest neighbour and kernel matching algorithms substantially improve balance over observable covariates between adopters and non-adopters within the sample.

Results indicate that adoption of improved groundnuts is associated with a two percentage point increase in household consumption expenditures, and a 12-17 percentage point reduction in

poverty headcount. The magnitude of estimated treatment effects increases for the poorest quartile of the sample, suggesting that improved groundnuts may be a pro-poor technology.

Critique: While PSM improves balance over observable covariates, no attempt is made to control for selection bias from unobservables, which include potentially important determinants of adoption such as ability, aspirations, and social capital. Thus, estimated treatments effects are potentially biased. Nonetheless, the use of PSM likely improves estimates relative to an endogenous OLS model.

2.3 Macro-level welfare analyses and returns on research investment
Tsusaka, Takuji, Harry Msere, Moses Siambi, Kizito Mazvimavi, and Patrick Okori, *Evolution and Impacts of Groundnut Research and Development in Malawi: An ex-post Analysis*, Vol. 11(3), pp. 139-158 (2016).

Summary: Despite considerable data limitations, Tsusaka et al. (2016) estimate the impacts of improved groundnut adoption and ROI for groundnut research in Malawi between 1983 (the year of initial investments) and 2013. Data are drawn from government and donor statistics, key informant interviews, and a non-random field survey of groundnut producers. Based on these data, the authors compute gains from improved groundnut adoption as the difference between moving averages of groundnut yields and labour inputs in the 1983-2013 period relative to the 1977-1982 period. Post-1983 gains are attributed to varietal improvement.

Comparison of 1983-2013 outcomes relative to those from 1977-1982 indicates that improved groundnut varieties are associated with yield increases of up to 1500 kg/ha and reductions in labour requirements of up to 30 days per year. Based on these comparative benefits, the internal rate of return for R&D investment in groundnut improvement is calculated at 22%, which is significantly higher than the prevailing interest rate in Malawi over the period (10.53%). This superiority is robust to reasonable changes in the interest rate, model choice, and assumed research or administrative costs of up to 50%, suggesting that R&D investments in groundnut improvements over 1983-2013 were both competitive and profitable.

Critique: Data availability and quality represent substantial challenges in this study, forcing the authors to estimate supply and demand elasticities, administrative and research costs, and improved groundnut yields. Variation of these estimates in sensitivity analysis demonstrates that results are relatively stable over alternative specifications, but does not correct bias from measurement error.

The larger problem concerns the use of outcomes from 1977-1982 as a counterfactual against which to measure impacts. The authors attribute post-1982 improvements in groundnut yields and labour requirements to the adoption of improved varieties, but improvements may just as well have resulted from alternative factors such as growth in farmers' skills, infrastructure, fertilizers, or other technologies – even in the absence of groundnut R&D. The data limitations

confronted by the authors make it impossible to identify the specific contribution of improved groundnuts to the observed gains, which in turn prevents reliable interpretation of the final ROI estimates.

Kumara Charyulu, D., D. Moses Shyam, Cynthia Bantilan, S.T. Borikar, S.K. Gupta, and K.N. Rai, *Pearl Millet Technology Adoption and Impact Study in Maharashtra*, Research Report 71, ICRISAT, 76 pp. (2016).

Summary: This study attempts to estimate unit cost reductions and total welfare benefits from the adoption of improved pearl millet varieties in Maharashtra, India. Estimates are based on archival data from ICRISAT and government sources, as well as a non-random survey of middle-aged, experienced farmers in Maharashtra. The authors divide pearl millet varieties into pre-2000 (unimproved) and post-2000 (improved) categories, and compute "impacts of improved variety adoption" as the difference in sample means unit costs between farmers growing improved varieties and farmers growing unimproved varieties. Results suggest that cultivation of improved varieties is associated with unit cost reductions of US$26.12 versus unimproved varieties, resulting in state-wide welfare gains of US$103 million.

Critique: As in the majority of macro-level welfare analyses, this study does not sufficiently focus on the valid identification of farm-level impacts. As explored in Part 2, the simple comparison of outcomes between adopters and non-adopters is likely to be biased by self-selection, which in turn biases all higher-level estimates based on this figure. Furthermore, the use of non-random sampling invalidates both internal and external validity.

Kumara, Charyulu D., D. Moses Shyam, Cynthia Bantilan, S.T. Borikar, A. Ashok Kumar, and Belum V.S. Reddy, *Rainy Season Sorghum Technology Adoption and Impact Study in Maharashtra*, Research Report 70, Patancheru 502 324, Telangana, India: ICRISAT (2016).

Summary: Kumara Charyulu et al. estimate average cost reductions and total welfare impacts associated with the adoption of improved rainy season sorghum varieties in Maharashtra between 1993 and 2022 (forecast). Estimates are computed based on data from a survey of 360 randomly selected sorghum farmers in Maharashtra in 2013. The sample is constructed from a selected sample of 20 sorghum-producing *tehsils* (a sub-district unit), with three villages randomly sampled within each *tehsil* and six households randomly sampled within each village. Sorghum varieties in the sample are categorized as improved (post-2000) or unimproved (pre-2000), and average cost reductions are computed as the sample mean difference between costs of production for farmers growing improved and unimproved varieties. This estimation methodology allows the authors to compute an average cost reduction of US$27 from adoption, and a total welfare gain of US$150 million between 1993-2022.

Critique: Analogous to the critiques outlined above, this study suffers from substantial selection bias due to its calculation of adoption impacts through a simple comparison of outcomes between adopters and non-adopters. This bias in the identification of farm-level impacts invalidates all higher-level estimates. Furthermore, a sample of size of 360 over an entire state,

and especially the sampling of only three villages per *tehsil* and six households per village, calls into question the representativeness and power of the sample (see Part 2.1).

Bantilan, C., D. Charyulu Kumara, P.M. Gaur, M.D. Shyam, and D. Jeff, *Short-Duration Chickpea Technology: Enabling Legumes Revolution in Andhra Pradesh, India*, Research Report No. 23. ICRISAT, 208 pp. (2014).

Summary: Bantilan et al. estimate farm-level unit cost impacts from the adoption of improved chickpea varieties in Andhra Pradesh, India, and input these estimates into an economic surplus model in order to estimate state-level welfare impacts from adoption. Farm-level estimates are based on a cross-sectional survey of 810 households randomly selected from 90 randomly selected villages within 30 randomly selected chickpea-producing districts in Andhra Pradesh. Of farmers in the sample, 98% had already adopted improved chickpea varieties, which mitigates risk of selection bias but creates the new problem of a missing counterfactual. The authors seek to create a counterfactual by eliciting farmers' estimates of what they *believe* their yields and costs would be if they were currently growing the old, unimproved chickpea variety. State-level welfare impacts are estimated based on data on R&D costs drawn from expert interviews. Based on this strategy, results indicate that adoption of improved chickpea varieties results in a US$144 per ton unit cost reduction at the farm-level, a US$359 million increase in total state welfare in 2012, and a ROI from R&D of US$146 million in 2014.

Critique: This study illustrates the methodological shortcomings often encountered in welfare analyses and ROI studies. To begin, the counterfactual estimate based on farmers' recall and guesswork is highly subject to present bias, enumerator influence, and subjective responses, not to mention severe endogeneity. As outlined in Part 2.4 above, valid identification of farm-level treatment effects is necessary to proceed with welfare analyses. Thus, this study's results are already invalid at this point. Furthermore, R&D expenses are estimated based solely on expert interviews, which (as outlined in Part 2.4) are often biased and inaccurate. Thus, even discounting the invalidity of farm-level impact estimates, estimates of ROI from R&D are likely to be highly inaccurate. In sum, results from this study offer little reliable information regarding the impacts of improved chickpea varieties in Andhra Pradesh.

Gierend, Albert, Henry Ojulong, and Nelson Wanyera, *A Combined Ex-Post/Ex-Ante Impact Analysis for Improved Sorghum and Finger Millet Varieties in Uganda*, ICRISAT, Socioeconomics Discussion Paper Series No. 19 (2014).

Summary: The authors of this study compile expert estimates on gains from adoption of improved sorghum and finger millet varieties, and costs of research and development in Uganda. These estimates are input into a DREAM economic surplus model in order to calculate ROI from R&D and total welfare impacts. Results from economic surplus modeling indicate that R&D for improved sorghum varieties yields total welfare benefits of US$125 million between 1980-2030, while R&D for improved finger millet varieties yields total welfare benefits of US$2.7 million

annually. Research ROI for sorghum is estimated at 28-60% over the 1980-2030 period, while ROI for R&D in finger millet is estimated to range between 96 and 107% over the same timeframe.

Critique: As established by Ndjeunga et al. (2013), expert estimates may diverge widely from values established in household surveys. Expert estimates may be especially biased when crop breeders and agricultural technicians are asked to estimate the gains from improved variety adoption, given that these individuals have substantial personal interest in the success of improved varieties. Thus, the results presented in this study are of low credibility due to reliance on expert estimates, and a failure to address the standard problem of selection bias.

Macharia, Ibrahim, Alastair Orr, Franklin Simtowe, and Solomon Asfaw, *Potential Economic and Poverty Impact of Improved Chickpea Technologies in Ethiopia*, International Association of Agricultural Economists Triennial Conference (2012).

Summary: This study seeks to measure the country-level welfare and poverty impacts of investments in R&D and promotion of improved chickpea varieties in Ethiopia between 1974 and 2005. Data on R&D and promotion costs are obtained from ICRISAT records; farm-level data are drawn from a two-wave panel survey conducted in Ethiopia in 2008 and 2010. However no further details regarding the survey are provided. The authors employ the DREAM model to estimate country-level impacts based on farm-level estimates of gains from adoption, which are computed as the gap between sample mean yields and returns for improved chickpea adopters versus non-adopters.

Survey findings indicate that improved varieties exhibit 33% higher yields and 31% higher profits than local varieties. Based on these values, the authors calculate that total welfare benefits from investments in improved chickpea R&D and promotion totaled US$111 million in 2005, constituting a 55% internal rate of return on investment. Based on estimates of the elasticity of poverty reduction with respect to growth in agricultural GDP, these gains from investment resulted in an estimated reduction in poverty of 7.4 million people in Ethiopia between 1974 and 2005 period. This accounting of costs and benefits does not account for the positive spillovers of agricultural R&D into other countries and sectors.

Critique: As emphasized in Part 2.4 above, the validity of country-level welfare and ROI analyses depends on the valid identification of farm-level impacts. By simply comparing outcomes for adopters versus non-adopters, this study fails to control for potential selection bias, likely inflating estimates of welfare gains. Therefore, results presented in this study are not reliable indicators of the true welfare and poverty impacts of investments in improved chickpea R&D and promotion in Ethiopia. It would be more productive to build a country-level welfare analysis of improved chickpea investments in Ethiopia upon the better-identified estimates of adoption impacts presented in Part 3.2.

Shiferaw, Bekele, Tewodros Kebede, and Liang You, Technology Adoption under Seed Access Constraints and the Economic Impacts of Improved Pigeonpea Varieties in Tanzania, *Agricultural Economics*, **Vol. 39, pp. 309-323 (2008).**

Summary: This study begins by estimating a double-hurdle model of improved pigeonpea adoption for a stratified random sample of 240 households in Babati, Tanzania in 2004. Efforts are made to demonstrate that Babati district is relatively balanced for relevant observable covariates with pigeonpea-producing regions of Tanzania as a whole. The double-hurdle model is a two-stage procedure, wherein the first stage estimates determinants of demand for improved pigeonpea and the second stage estimates determinants of access to improved seed, so that adoption occurs only if demand is positive and seed access is achieved.

After controlling for observables, adoption is found to be correlated with 67% higher yields and 46% higher net returns, with gains largely attributable to improved varieties' disease resistance. Next, these farm-level gains are extrapolated to the country level using the DREAM economic surplus model and national statistics on rates of adoption and costs of pigeonpea R&D. Results indicate that investment in pigeonpea R&D yielded a 32.2% ROI and total welfare benefits of US$6.1 million in 2004.

Critique: This study constitutes a step in the right direction in terms of methodological approaches to country-level welfare analyses, insofar as farm-level gains from adoption are estimated conscientiously through a modeling procedure. Nonetheless, substantial problems persist in this study, calling the reliability of the results into question. First, the study is based on a small sample from only one district, and the balance between Babati and Tanzania as a whole is only demonstrated over a small set of observables. It is reasonable to suspect that Babati differs from the rest of Tanzania over a range of important unobserved variables, calling into question the external validity of the findings and their applicability to country-level estimates.

Second, while the work of Shiferaw et al. demonstrates a significant advance over simple correlation studies by controlling for a wide range of observable differences between adopters and non-adopters, their findings are still subject to selection bias over unobservables. Methods such as PSM or IV would allow for more convincing identification of the causal effects of adoption in this setting.

2.4 Correlation studies and other methods
Kaliba, A, K. Mazvimavi, and G. Ghebreyesus, Economic Profitability and Risk Analyses of Improved Sorghum Varieties in Tanzania, *Journal of Development and Agricultural Economics*, **Vol. 9, No. 9, pp. 250-268 (2017).**

Summary: Kaliba et al. compare distributions of net returns for both improved and local sorghum varieties in Tanzania. Analyses are based on data collected from village-level focus groups (with total participation of 822 households) in Central, Western, and Northern Tanzania during 2012-2013. Data on planted area and returns, agricultural practices, and socioeconomic variables are

used to estimate distributions of net returns for different combinations of varieties and practices. Results indicate that net returns from improved sorghum varieties stochastically dominate local varieties. In other words, improved varieties offer higher net returns without higher volatility.

Critique: While this study offers an insightful analysis of the interactions and risk-tradeoffs between different sorghum varieties and agronomic practices, it does not employ econometric methods to identify the causal effect of adoption. Therefore, the observed mean and variance shifts in net returns between improved and local varieties are endogenous correlations, since it is unlikely that farmers' choices about crops and agronomic practices are independent of other factors determining yields and returns. Further consideration should be made of how to control for potential simultaneity and selection bias from omitted variables.

Kumara, Charyulu D, Cynthia Bantilan, A. Rajalaxmi, B.V.S. Reddy, S.T. Borikar, Kumar A. Ashok, N.P. Singh and Shyam D. Moses, ***Development and Diffusion of Sorghum Improved Cultivars in India: Impact on Growth and Stability in Yield***. **Working Paper Series No. 50. 92 pp. (2014).**

Summary: This study is primarily focused on providing a historical overview of sorghum research and policy in India. Data are drawn from ICRISAT archives and the Indian Ministry of Agriculture. In the section of the study relevant to impact assessment, regression analysis is employed to explore associations between rates of improved variety adoption, yields, yield variance, and control variables at the district level. Regressions reveal that higher adoption rates are associated with higher yields, but that adoption rates do not have a significant association with yield variance.

Critique: The study provides valuable historical information on improved sorghum research and diffusion in India. Nonetheless, the positive correlation between higher adoption rates and higher district-level yields should not be taken as evidence that adoption causes yield gains for the standard reasons: potential endogeneity from simultaneity, omitted variables, and measurement error.

Karunakaran K.R., C.P. Gracy, H. Lokesha, C. Bantilan, D. Kumara Charyulu, Rao P. Parthasarathy, G.D. Nageswara Rao, M. Vaithiyalingan, H.L. Nadaf, P. Venkataramana, H.D. Upadhyaya, P. Janila, and K.P.C. Rao. ***Groundnut Baseline and Early-Adoption Surveys in South Asia: Insights from TL-II (Phase-1) Project***: ***Synthesis Report,*** **ICRISAT, 104 pp (2013).**

Summary: This study is focused more on adoption and farmer participatory variety selection (FPVS) programmes than on measuring the impacts of improved varieties. Data are drawn from an experimental trial of FPVS programmes in Karnataka and Tamil Nadu, India. FPVS programmes are administered in treatment villages but not in control villages, and adoption rates and yields from baseline and follow up surveys are compared. FPVS "treatment" is associated with higher

rates of adoption of improved varieties, and improved varieties are found to exhibit higher net returns than local varieties.

Critique: Identification of average treatment effects from FPVS requires random assignment of treatment. Nonetheless, it is not clear that treatment is assigned randomly either at the village or household level. At the village level, pre-treatment balance is not effectively demonstrated, while at the individual level, measures are not taken to avoid selection-into-treatment. Thus, the FPVS trial described in this study is only experimental in its general design, not in implementation. Furthermore, the positive association between improved variety adoption and higher net returns is potentially biased by self-selection into adoption.

Nwankwo Chianu, Jonas, Jeroen Huising, Seth Danso, Peter Okoth, Justina Nwanganga Chianu, and Nteranya Sanginga, Financial Value of Nitrogen Fixation in Soybean in Africa: Increasing Benefits for Smallholder Farmers, *Journal of Life Sciences***, Vol. 4, No. 31 (2010).**

Summary: Nitrogen fixation in soils is one of the primary environmental benefits of legume cultivation. Nitrogen fixation reduces the need for chemical fertilizers and enhances the productivity of both legume-cereal intercropping and rotation systems, potentially increasing yields and incomes. and reducing chemical leaching and runoff into groundwater.

Nwankwo et al. estimate the financial savings (in cost-equivalent units of fertilizer) from soybean nitrogen fixation in sub-Saharan Africa. Data on nitrogen fixation levels for different soybean varieties in both rotation and intercropping systems, as well as data on nutrient use efficiency in complementary cereal crops, are drawn from a survey of the technical literature. Data on fertilizer prices are drawn from a report on the global fertilizer market. Time series on harvested soybean area for a group of African countries are drawn from FAOSTAT.

Findings indicate that in Nigeria, a 10% increase in area planted with soybeans would result in N-fixation worth US$44 million annually, as well as a 20% increase in cereal yields. Furthermore, the authors calculate that total soybean cultivation in a group of 19 African countries results in a urea-cost-equivalent savings of US$203 million during 2002-2004. The majority of impact is concentrated in Nigeria and South Africa due to the dominance of African soybean cultivation. Across all developing countries, the authors estimate the value of N-fixation from legume crops at US$6.7 billion per year.

Critique: Results are based on data drawn from the technical literature on leguminous N-fixation, which may or may not accurately reflect the realities of smallholder production in sub-Saharan Africa. In addition, theoretical cost savings from fertilizer substitution is of limited relevance in contexts where farmers overwhelmingly do not purchase chemical fertilizers. Nonetheless, the study makes a valuable contribution by illustrating the scale of potential environmental benefits from legume cultivation, especially in environmentally degraded agro-ecologies such as those in sub-Saharan Africa.

Appendix 3: Summary tables

Figure 3. Impact evaluation themes for GLDC Crops

		Directly evaluated impacts				
Study	**Crop**	**Micro-economic**	**Macro-economic**	**Environ-mental**	**Nutrition & food security**	**Social**
Michler et al., 2018	Chickpea	▨				
Verkaart et al., 2017	Chickpea	▨				
Bantilan et al., 2014	Chickpea		▨			
Macharia et al., 2012	Chickpea		▨			
Asfaw et al., 2012	Chickpea	▨				
Tsusaka et al., 2016	Groundnut		▨			
Ndjeunga et al., 2013	Groundnut	▨				
Karunakaran et al., 2013	Groundnut					
Simtowe et al., 2012	Groundnut	▨				
Kumara Charyulu et al., 2016	Pearl millet		▨			
Finkelstein et al., 2015	Pearl millet				▨	
Kaliba et al., 2017	Sorghum	▨				
Kumara Charyulu et al., 2016	Sorghum		▨			
Kumara Charyulu et al., 2014	Sorghum		▨			
Gierend et al., 2014	Sorghum		▨			
Asfaw et al., 2012	Pigeonpea	▨				
Shiferaw et al., 2008	Pigeonpea	▨	▨			
Bulte et al., 2014	Cowpea	▨				
Nwankwo et al., 2010	Soybean			▨		
Gierend et al., 2014	Finger millet		▨			
Total		9	9	1	1	0

Study	Crop	Country/region	Intervention	External validity	Estimated impact	Methodology
Bulte et al., 2014	Cowpea	Tanzania	Improved seed treatment in RCT	Cowpea producers in Morongoro, Tanzania	27% increase in harvest volume from improved variety treatment, but 2/3 of effect due to induced effort (placebo).	Experimental (RCT and double-blind RCT)
Finkelstein et al., 2015	Pearl Millet	India	Inclusion of biofortified pearl millet in school lunch programme	Upper bound of potential treatment effects on school children in India	Statistically significant positive treatment effects on four measures of iron deficiency	Experimental (double-blind RCT)
Michler et al., 2018	Chickpea	Ethiopia	Adoption of improved varieties	Chickpea-producing regions in Ethiopia	Adoption increases household profits by 23-28%	Observational: fixed effects, correlated random coefficients
Verkaart et al., 2017	Chickpea	Ethiopia	Adoption of improved varieties	Chickpea-producing regions in Ethiopia	Adoption significantly increases household income and reduces probability of poverty	Observational: IV and multi-hurdle
Ndjeunga et al., 2013	Groundnut	Nigeria	Introduction of improved varieties	Semi-arid groundnut-producing regions of Nigeria	Adoption significantly increases yield and consumption, but only increases wealth for rich farmers	Observational: IV and PSM
Asfaw et al., 2012	Chickpea	Ethiopia	Introduction of improved varieties	Chickpea-producing regions in Ethiopia	Adoption associated with 24.6% increase in welfare for adopters, 20.9% for non-adopters	Observational: IV, endogenous switching regression
Asfaw et al., 2012	Pigeonpea	Tanzania	Introduction of improved varieties	Pigeonpea-producing regions of Tanzania	Adoption associated with 100.3% increase in welfare for adopters & 99.4% for non-adopters	Observational: IV, endogenous switching regression

Simtowe et al., 2012	Groundnut	Malawi	Adoption of improved varieties	Groundnut-producing regions of Malawi	Adoption associated with 2% increase in consumption and 12-17% decrease in poverty	Observational: PSM

Figure 5. Summary of GLDC impact evaluations

Figure 6. Summary of GLDC impact evaluations (continued)

Study	Crop	Country/region	Intervention	External validity	Estimated impact	Methodology
Tsusaka et al., 2016	Groundnut	Malawi	Research and development of improved varieties	No external validity	Research ROI of 22%, compared to interest rate of 10%	Ex-post economic surplus model (welfare and ROI)
Kumara Charyulu et al., 2016	Pearl millet	Maharashtra, India	Introduction of improved varieties	No external validity	Welfare gain from adoption of US$103 million from 1993-2022	Descriptive statistics; ex-post economic surplus model
Kumara Charyulu et al., 2016	Sorghum	Maharashtra, India	Introduction of improved varieties	Sorghum-producing regions of Maharashtra	Welfare gain from adoption of US$150 million between 1993-2022	Ex-post economic surplus model (welfare and ROI)
Bantilan et al., 2014	Chickpea	Andhra Pradesh, India	Introduction of improved varieties	Chickpea-producing regions of Andhra Pradesh	Welfare gain from adoption of US$359 million in 2012 & research ROI of USD$146 million in 2014	Ex-post economic surplus model (welfare and ROI)
Gierend et al., 2014	Sorghum	Uganda	Introduction of improved varieties	No external validity (expert estimates)	Welfare gain from adoption of US$125 million between 1980 and 2030, & research ROI of 28-60%	Ex-post economic surplus model (welfare and ROI)
Gierend et al., 2014	Finger millet	Uganda	Introduction of improved varieties	No external validity (expert estimates)	Welfare gain from adoption of USD$2.7 million annually; Research ROI of 96% and 107%	Ex-post economic surplus model (welfare and ROI)
Macharia et al., 2012	Chickpea	Ethiopia	Research and development of improved varieties	No external validity	Welfare gain from adoption of US$111 million (1974-2005 & research ROI of 5:1	Ex-post Economic Surplus Model (welfare and ROI)

| Shiferaw et al., 2008 | Pigeonpea | Tanzania | Introduction of improved varieties | Pigeonpea-producing areas of Babati District, Tanzania | Welfare gains from adoption of US$6.1 million in 2004, and research ROI of 32.2% | Double-hurdle model of adoption; ex-post economic surplus model |

Figure 7. Summary of GLDC impact evaluations (continued)

Study	Crop	Country/region	Intervention	External validity	Estimated impact	Methodology
Kaliba et al., 2017	Sorghum	Tanzania	Adoption of improved varieties	Sorghum-producing regions of Central, Western, and Northern Tanzania	Improved varieties exhibit significantly higher yields and profits than local varieties	Descriptive statistics/cost-benefit analysis
Kumara Charyulu et al., 2014	Sorghum	India	Introduction of improved varieties	No external validity (expert estimates)	Productivity levels have risen and yield variability has declined due to adoption of improved varieties	Descriptive statistics
Karunakaran et al., 2013	Groundnut	Southern India	Farmer participatory variety selection programmes	No external validity	Improved varieties exhibit higher net returns than local varieties	Descriptive statistics
Nwankwo et al., 2010	Soybean	Sub-Saharan Africa	Introduction of improved varieties	Model-based; no survey data	Savings of US$167/ha from soybean cultivation, total savings of US$203 million between 2002-2004 over 19 African countries	Cost reduction model

Working Paper Series

240. The national agroforestry policy of India: experiential learning in development and delivery phases. http://dx.doi.org/10.5716/WP16143.PDF

241. Agroforestry and forestry in Sulawesi series: Livelihood strategies and land-use system dynamics in Gorontalo. http://dx.doi.org/10.5716/WP16157.PDF

242. Seri Agroforestri dan Kehutanan di Sulawesi: Strategi mata pencaharian dan dinamika system penggunaan lahan di Gorontalo. http://dx.doi.org/10.5716/WP16158.PDF

243. Ruang, Gender dan Kualitas Hidup Manusia: Sebuah studi Gender pada komunitas perantau dan pengelola kebun di Jawa Barat. http://dx.doi.org/10.5716/WP16159.PDF

244. Gendered knowledge and perception in managing grassland areas in East Sumba, Indonesia. http://dx.doi.org/10.5716/WP16160.PDF

245. Pengetahuan dan persepsi masyarakat pengelola padang aavana, Sebuah Kajian Gender di Sumba Timur. http://dx.doi.org/10.5716/WP16161.PDF

246. Dinamika Pengambilan Keputusan pada komunitas perantau dan pengelola kebun di Jawa Barat. http://dx.doi.org/10.5716/WP16162.PDF

247. Gaharu (eaglewood) domestication: biotechnology, markets and agroforestry options. http://dx.doi.org/10.5716/WP16163.PDF

248. Marine habitats of the Lamu-Kiunga coast: an assessment of biodiversity value, threats and opportunities. http://dx.doi.org/10.5716/WP16167.PDF

249. Assessment of the biodiversity in terrestrial landscapes of the Witu protected area and surroundings, Lamu County, Kenya. http://dx.doi.org/10.5716/WP16172.PDF

250. An ecosystem services perspective on benefits that people derive from biodiversity of Coastal forests in Lamu County, Kenya http://dx.doi.org/10.5716/WP16173.PDF

251. Assessment of the biodiversity in terrestrial and marine landscapes of the proposed Laga Badana National Park and surrounding areas, Jubaland, Somalia.

http://dx.doi.org/10.5716/WP16174.PDF

2017

252. Preferensi Petani terhadap Topik Penyuluhan dan Penyebaran Informasi Agroforestri di Indonesia   http://dx.doi.org/10.5716/WP16181.PDF

253. Seri Agroforestri dan Kehutanan di Sulawesi: Keanekaragaman hayati jenis pohon pada hutan rakyat agroforestri di DAS Balangtieng, Sulawesi Selatan
http://dx.doi.org/10.5716/WP16182.PDF

254. Potensi dan Tantangan dalam Pengembangan Skema Ko-Investasi Jasa Lingkungan di Kabupaten Buol, Indonesia. http://dx.doi.org/10.5716/WP17008.PDF

255. Keragaman Jenis Pohon dan Pemanfaatannya oleh Masyarakat di Kabupaten Buol, Indonesia. http://dx.doi.org/10.5716/WP17009.PDF

256. Kerentanan dan preferensi sistem pertanian petani di Kabupaten Buol, Indonesia
http://dx.doi.org/10.5716/WP17010.PDF

257. Dinamika Perubahan Penggunaan/Tutupan Lahan Serta Cadangan Karbon di Kabupaten Buol, Indonesia. http://dx.doi.org/10.5716/WP17011.PDF

258. The effectiveness of the volunteer farmer trainer approach vis-à-vis other information sources in dissemination of livestock feed technologies in Uganda.
http://dx.doi.org/10.5716/WP17104.PDF

259. Agroforestry and forestry in Sulawesi series: Impact of agricultural-extension booklets on community livelihoods in South and Southeast Sulawesi.
http://dx.doi.org/10.5716/WP17125.PDF

260. Petani Menjadi Penyuluh, Mungkinkah? Sebuah Pendekatan Penyuluhan dari Petani ke Petani di Kabupaten Sumb Timur. http://dx.doi.org/10.5716/WP17145.PDF

261. Dampak Perubahan Tutupan Lahan terhadap Kondisi Hidrologi di Das Buol, Kabupaten Buol, Sulawesi Tengah: Simulasi dengan Model Genriver
http://dx.doi.org/10.5716/WP17146.PDF

262. Analisis Tapak Mata Air Umbulan, Pasuruan, Jawa Timur. Kajian elemen biofisik dan persepsi masyarakat.   http://dx.doi.org/10.5716/WP17147.PDF

263. Planned comparisons demystified. http://dx.doi.org/10.5716/WP17354.PDF

264. Soil health decision support for NERC digital soil platforms: A survey report
http://dx.doi.org/10.5716/WP17355.PDF

265. Seri Pembangunan Ekonomi Pedesaan Indonesia: Menanam di bukit gundul:
Pengetahuan masyarakat lokal dalam upaya restorasi lahan di Sumba
Timur. http://dx.doi.org/10.5716/WP17356.PDF

266. Tree diversity and carbon stock in three districts of Kutai Timur, Pasir and Berau, East
Kalimantan http://dx.doi.org/10.5716/WP17357.PDF

267. Tree Diversity and Carbon Stock in Various Land Use Systems of Banyuasin and Musi
Banyuasin Districts, South Sumatera http://dx.doi.org/10.5716/WP17358.PDF

268. Tree diversity and carbon stock in various land cover systems of Jayapura, Jayawijaya
and Merauke Districts, Papua Province http://dx.doi.org/10.5716/WP17359.PDF

269. Modelling tree production based on farmers' knowledge: case for kapok (Ceiba
pentandra) and candlenut (Aleurites mollucana) under various agroforestry scenarios.
http://dx.doi.org/10.5716/WP17361.PDF

270. The Impact of Land Cover and Climate Change on Present and Future Watershed
Condition. Study case: Tugasan, Alanib and Kulasihan Sub-watershed of Manupali
Watershed, Lantapan, Bukidnon, Philippines. http://dx.doi.org/10.5716/WP17362.PDF

271. Tree Diversity and Above-ground Carbon Stock estimation in Various Land use Systems
in Banjarnegara, Banyumas and Purbalingga, Central Java.
http://dx.doi.org/10.5716/WP17363.PDF

272. Agroforestry and Forestry in Sulawesi series: Landscape Management Strategies in
Sulawesi: Review of Intervention Options. http://dx.doi.org/10.5716/WP17364.PDF

273. Household Food-Security and Nutritional Status of Women and Children in Buol
Regency, Central Sulawesi, Indonesia. http://dx.doi.org/10.5716/WP17365.PDF

274. Palm oil expansion in tropical forest margins or sustainability of production? Focal
issues of regulations and private standards. http://dx.doi.org/10.5716/WP17366.PDF

2018

275. Decision analysis methods guide: Agricultural policy for nutrition
http://dx.doi.org/10.5716/WP18001.PDF

276. Supporting human nutrition in Africa through the integration of new and orphan crops into food systems: Placing the work of the African Orphan Crops Consortium in context. http://dx.doi.org/10.5716/WP18003.PDF

277. Seri Pembangunan Ekonomi Pedesaan Indonesia. Pilihan Manajemen Budidaya Kacang Tanah sebagai Upaya untuk Memperbaiki Penghidupan Masyarakat Haharu. http://dx.doi.org/10.5716/WP18004.PDF

278. Estudio de línea de base CCAFS a nivel de hogar en Nicaragua y Costa Rica Fase de diagnóstico del estudio: "Contribución de la diversidad arbórea a los medios de vida para la adaptación y la mitigación al cambio climático http://dx.doi.org/10.5716/WP18005.PDF

279. Understanding tree cover transition, drivers and stakeholder perspectives for effective landscape governance. A case study in Na Nhan commune, Dien Bien province, Vietnam. http://dx.doi.org/10.5716/WP18006.PDF

280. El Sistema "Quesungual": Agroforestería y manejo de suelos para la producción de maíz y frijol en laderas. http://dx.doi.org/10.5716/WP18007.PDF

281: Probabilistic Decision Modelling to Determine Impacts on Natural Resource Management and Livelihood Resilience in Marsabit County, Kenya.

http://dx.doi.org/10.5716/WP18008.PDF

282. Shifting discourse, shifting power: how is climate change mitigation and justice negotiated in Indonesia? http://dx.doi.org/10.5716/WP18009.PDF

283. Result of Land Use Planning and Land Administration (LULA) Implementation in South Sumatra, East Kalimantan, Central Java and Papua http://dx.doi.org/10.5716/WP18010.PDF

284. Farmers' preferences for training topics and dissemination of agroforestry information in Indonesia. http://dx.doi.org/10.5716/WP18015.PDF

285. CSA-Diagnostic (CSA-Dx): A primer for investigating the 'climate-smartness' of ag technologies http://dx.doi.org/10.5716/WP18020.PDF

286. An analysis of the vulnerability of poor communities in Yunnan Province, China

http://dx.doi.org/10.5716/WP18021.PDF

287. Gendered space and quality of life: gender study of out-migration and smallholding agroforestry communities in West Java Province, Indonesia. http://dx.doi.org/10.5716/WP18024.PDF

288: Evaluation of UTZ certification coffee businesses in Guatemala, Honduras and Nicaragua. http://dx.doi.org/10.5716/WP18028.PDF

289. Agroforestry species of Peru: annotated list and contribution to prioritization for genetic conservation. http://dx.doi.org/10.5716/WP18029.PDF

290. Indonesia Rural Economic Development Series.Growing plants on a barren hill: local knowledge as part of land restoration in Sumba Timur, Indonesia.
http://dx.doi.org/10.5716/WP18030.PDF

291. Assessing the Downstream Socioeconomic Impacts of Agroforestry in Kenya

http://dx.doi.org/10.5716/WP18033.PDF

2019

292: Los árboles fuera del bosque en la NAMA forestal de Colombia. Elementos conceptuales para su contabilización. http://dx.doi.org/10.5716/WP19002.PDF

293: Gender and Adaptation: An Analysis of Poverty and Vulnerability in Yunnan, China. DOI: http://dx.doi.org/10.5716/WP19004.PDF

294: Tree Cover on Agricultural Land in the Asia Pacific Region. DOI: http://dx.doi.org/10.5716/WP19005.PDF

World Agroforestry (ICRAF) is a centre of scientific and development excellence that harnesses the benefits of trees for people and the environment. Leveraging the world's largest repository of agroforestry science and information, we develop knowledge practices, from farmers' fields to the global sphere, to ensure food security and environmental sustainability.

ICRAF is the only institution that does globally significant agroforestry research in and for all of the developing tropics. Knowledge produced by ICRAF enables governments, development agencies and farmers to utilize the power of trees to make farming and livelihoods more environmentally, socially and economically sustainable at multiple scales.

World Agroforestry