Rivista Internazionale di Scienze Sociali, 2015, n. 4, pp. 369-380

THE ROLE OF IMPACT EVALUATION OF AGRICULTURAL DEVELOPMENT PROJECTS

Jacopo Bonan*
Stefano Pareglio**
Valentina Rotondi***

ABSTRACT

Agricultural development is central in the debate on the post-2015 development agenda, food security and poverty eradication being directly linked to agricultural production. Development actors are working to translate identified priorities into implementable policies aimed at increasing small farmers' productivity, which is one of the most important barriers to agricultural development. Agricultural extension is deemed as a way to transfer innovation to farmers. Despite the development of extension programs, few studies investigate their impacts through rigorous evaluation methods. Impact evaluation exercises, by incorporating the methodology of the counterfactual, play a pivotal role in identifying the effects of agricultural programs in a causal manner, allowing the selection of the most cost-effective interventions and providing information for evidence-based policy making. The paper underlines the role of impact evaluation methods applied to agricultural development programs. It focuses on agricultural extension projects and briefly presents the results of a case study in rural Ethiopia.

Keywords: Agricultural Extension, Rural Development, Impact Evaluation, Food Security, Diet Diversification.
JEL Classification: Q10, Q18, D04.

1. INTRODUCTION

A large proportion of the world's poor live in rural areas (FAO, 2014) and for 70% of them agriculture is the main source of income and employment (The World Bank, 2015). Inclusive growth, poverty alleviation and food security are directly linked to agriculture, which is one of the main focuses of the debate on the new post-2015 development agenda. In fact, agriculture plays a pivotal role in nur-

We acknowledge the financial support of Fondazione punto.sud and Fondazione Giangiacomo Feltrinelli. We are very thankful to CIAI.

* Jacopo Bonan, Department of Mathematics and Physics, Catholic University of the Sacred Heart, LabExpo (Feltrinelli Foundation) and FEEM. Email: jacopodaniele.bonan@unicatt.it.

** Stefano Pareglio, Department of Mathematics and physics, Catholic University of the Sacred Heart. Email: stefano.pareglio@unicatt.it.

*** Valentina Rotondi, Department of Economics, Catholic University of the Sacred Heart. Email: valentina.rotondi@unicatt.it.

turing healthy ecosystems and in supporting the sustainable management of land, water and natural resources, while ensuring world food security.

In order to be effective in fostering a dynamic and inclusive growth, agriculture requires a productivity revolution in smallholder farming (Christiaensen and Demery, 2007). One important barrier to the increase in farmers' productivity is linked to informational constraints, low diffusion of basic agricultural knowledge and slow adoption of good practices and simple technologies. Several countries have tried to fill this gap through agricultural extension programs aimed at transferring information from both the global knowledge base and local research to farmers, thus enabling them to clarify their own goals and possibilities, helping them to make better decisions, and stimulating desirable agricultural development (van der Ban and Hawkins, 1996). Despite the effort put into these programs all over the world[1], in sub-Saharan Africa agriculture is still at subsistence level and technology adoption is still scarce (Krishnan and Patnam, 2014). Understanding how new technologies and good practices are (not) adopted by farmers is therefore an open issue and has been an important focus of the agricultural economics research agenda (Udry, 2010). Relatively few studies attempt to shed light on the actual impacts of agricultural development programs in general and on agricultural extension in particular (DFID, 2014). For this purpose, the implementation of monitoring and evaluation tools is increasing at all levels, with the aim to provide impact assessments. These can be used by implementers to improve the quality, efficiency and cost-effectiveness of interventions, as the focus is on the results, rather than on the inputs and outputs. Moreover, measuring program impacts raises public awareness and promotes accountability by stakeholders, while, at a more global level, it contributes to building knowledge on what does and does not work to reduce poverty, improve welfare and sustain development.

In this paper we briefly review impact evaluation methodologies which make use of the counterfactual, and we describe how they have been applied to agricultural extension in developing countries. We then present a case study in rural Ethiopia. The aim of this work is to contribute to the debate on the importance of rigorously assessing the real impacts of development programs in order to influence policy making by informing all development actors: donors, implementers, governments and communities. Empirical methods using the counterfactual are now becoming a standard approach for many institutions and international organizations, which are now endowed with specific research divisions dedicated to impact evaluation. Several foundations and international research centers are currently dedicating relevant investments to the evaluation of programs. We think that a better understanding of this methodology and its systematic application to the evalua-

[1] Despite the dramatic decline in budgetary support for extension services in many developing countries over the past 20 years, large resources are still dedicated to such activities, for example the Ethiopian government has dedicated resources close to 1% GDP on extension services, in the last decade.

tion of development projects will be a key tool in adopting effective development strategies.

The rest of this paper is organized as follows: Section 2 presents the impact evaluation methodology. Section 3 reviews the literature on the evaluation of extension programs in agriculture while focusing in particular on rigorous evaluations. Section 4 presents a case study on rural Ethiopia. Section 5 concludes and provides some policy recommendations.

2. IMPACT EVALUATION METHODS

Monitoring and impact evaluation systems constitute the backbone of evidence-based public policies and programs. Impact evaluation investigates the change in outcomes of policy relevance that are directly attributable to the program, in other words it seeks causal effects. It shows what works and what does not, suggesting what programs are worth scaling up. Considerations on the cost-effectiveness and efficiency of policies and programs based on the evidence collected through impact evaluation studies are extremely important when resources are scarce, as they allow identifying the best way to spend money for the goals of interest. Several methods have been proposed to provide robust evidence to policy makers on the impact of development programs.

In general terms, a rigorous impact evaluation requires the identification of the counterfactual in order to answer the question on what would have happened if the intervention had not been in place. There is a fundamental difference between monitoring outcomes, which is limited to the description of the factual (what happens to project beneficiaries) and assessing impacts with the counterfactual, which allows attributing changes in the outcomes to the intervention in a causal manner.

Assessing the impact of agricultural development policies, in our case extension services, requires addressing the attribution issue through the use of identifications which tackle endogeneity and selection bias problems. Endogeneity arises, for instance, from the fact that allocation of extension efforts may not necessarily be random across or within localities. For example, targeting more advantaged (disadvantaged) areas with extension services may respond to different policy aims, but would bias impact estimations upwards (downwards). In other words, if we compared advantaged targeted localities with disadvantaged untargeted ones we would be unable to disentangle the actual impact of the project from the other factors such as the higher quality of soils, better connectedness to markets, infrastructure development, etc. The consequence of this would necessarily be an overestimation of the project's impacts.

Selection bias problems occur when observable and unobservable differences between treated and non-treated individuals cannot be distinguished from the true project impact, as the project's process of selection into the project is driven by observable or unobservable characteristics which influence the outcome. For example, in the case of extension services, some farmers may decide to participate in

project activities whereas others may decide not to do so. This decision is unlikely to be random, instead, it is related to personal characteristics such as the level of pro-activity, curiosity, propensity to learn and unobserved ability. Once again, the comparison of the outcomes of participants and non-participants would lead to misconclusions on the project's impacts: participants are likely to perform better in agriculture than non-participants even in the absence of project activities. The difference in outcomes would thus incorporate both the impact of the project and the role of personal characteristics which led to both project participation and better agricultural outcomes. Once again this would lead to an overestimation of the project's impact.

In order to correctly assess the causal impact of an agricultural extension project, without including misleading factors, one would need to compare, for the same subject, what happened under project participation (the factual) with what would have happened without the program (the counterfactual). However, this exercise is never possible. Different statistical methods have been suggested to identify counterfactuals through control groups, allowing the proper comparison of outcomes and leading to correct cause-effect assessments.

Impact evaluation methods using the counterfactual can be classified in two main categories: experimental and non-experimental designs. Experimental designs construct the counterfactual through the random assignment of individuals to either treatment or control groups. As a consequence of that, individuals from the two groups differ only in being (randomly) exposed to the program or not. This allows the attribution of changes in the outcomes of interest to the intervention in a causal manner. Randomized Controlled Trials (RCT), originally used in medical and epidemiological studies, have become an important methodology in development microeconomics over the last decade (Banerjee and Duflo, 2012). Non-experimental methods are used when it is not possible to directly manipulate the assignment process of project exposure and rely on observational data, derived from the observation of the natural development of events. The identification of the counterfactual using non-experimental methods requires specific, fairly strong assumptions. The most common techniques are difference-in-differences (DID), propensity score matching (PSP), regression discontinuity design (RDD), instrumental variables (IV) and panel data. All methods have important limitations and pitfalls which should be taken into consideration[2].

The implementation of impact evaluation systems is not exempt from limits and critical issues. First of all, costs: evaluation designs guaranteeing sound inferences require large samples of both treated and control units that are administered long questionnaires, possibly in several points in time. This is more expensive than collecting qualitative data or performance indices. Second, impact evaluation analysis like RCT may raise ethical issues, when the provision of benefits from pro-

[2] For an exhaustive discussion of experimental and non-experimental methods, their limits and potentials see Imbens and Wooldridge (2009).

jects or programs is only attributed to the treatment group, while the control group is excluded. Although several operative solutions have been suggested[3], further reflection is needed, should impact evaluation analysis become a standard and common procedure for policies and programs. Great attention to the ''*do no harm*'' principle and to informed consent should be devoted to all experimental circumstances. Third, the issue of generalization, or external validity: a policy, which may be successful in a given place, may not necessarily work elsewhere in the same way. Evaluations should therefore go beyond the basic assessment of impacts, trying to investigate the mechanisms and processes at work, aiming at more generalizable conclusions. Replicating impact evaluations of projects in different geographical contexts is a key practice in order to draw general lessons. Regular systematic reviews of the available literature on the impact of particular interventions replicated in different geographical areas, for instance through the use of meta-analyses, are of extreme importance in order to draw generalized conclusions.

3. IMPACT EVALUATION OF AGRICULTURAL EXTENSION

Agricultural extension encompasses the entire set of organizations and activities that support and facilitate people engaged in agricultural production to solve problems and to obtain information, skills, and technologies to improve their livelihoods, well-being and productivity (Birner et al., 2009). Whilst there is demand for evidence of the effectiveness of these programs, the empirical literature often gives mixed results. Several scholars show that agricultural extension has a positive effect on several rural livelihoods including knowledge, skills, productivity, consumption, and poverty reduction (e.g. Dercon et al., 2009) while others find only negligible achievements (Rivera et al., 2001). These findings differ according to the technology and context of project implementation, while the estimation of the impact of these programs is often hampered by attribution issues, endogeneity and selection bias. The causal interpretation of the results can therefore be challenging. In the field of agricultural extension, few rigorous impact evaluations explicitly tackle the issue of endogeneity (Aker, 2011) through different empirical strategies. In what follows we present the main findings on extension impacts, obtained through different methodologies, and we shortly describe the main insights of each.

Dercon et al. (2009) make use of an IV approach to assess the impact of public investments in agricultural extension and road quality on consumption growth and

[3] Among the possible solutions, the phase-in-design exploits different timings in the assignment of the treatment to obtain the control group which is not excluded from the intervention; the encouragement design allows to random vary incentives to be treated or not, without preventing anybody from the intervention; oversubscription and public lotteries allow to randomly identify treatment and control group in a group of eligible beneficiaries, in a transparent and fair way.

poverty rates in rural Ethiopia. Impact measures using IV are based on the identification of a proper, informative instrumental variable, i.e. strongly correlated to the endogenous variable (receiving the treatment) and valid, i.e. unrelated to the outcome of interest (consumption growth and poverty rates). The authors find that receiving at least one extension visit reduces headcount poverty by 9,8% and increases consumption growth by 7,1%. The identifying assumption relies on the fact that variations driven by the instrument on the outcome variable should occur only through the instrumented variable, i.e. the endogenous factor.

Godtland et al. (2004) estimate the effect of a particular form of agricultural extension service, delivered through farmer field schools, on farmers' knowledge of integrated pest management practices. In order to tackle selection bias issues, they use propensity score matching (PSM) methods to build a statistical comparison group of farmers comparable to FFS graduates. Matching is a non-experimental technique that allows the evaluation of treatment effect by comparing treated (FFS participants) and non-treated (non-participants) units, which are matched along a set of several observable characteristics summarized in a scalar, called propensity score. They find that farmers who participate in the program have significantly more knowledge about integrated pest management practices than those in the non-participant comparison group. The identifying assumption is that, upon matching the most relevant observables, the two samples are random, although the method is unable to account for unobserved heterogeneity.

Davis et al. (2012) also provide evidence of the impacts of a farmer field school project in East Africa on the economic and production spheres by using a combination of DID and matching estimation. DID design is based on the observations of, respectively, groups of households living in villages where FFS activities were implemented (treatment group) or not implemented (control group)[4]. Surveys carried out at two different times, before and after the program's development, allow us to identify the program's effects by calculating the double difference between the outcomes in the treatment vs control groups both before and after program implementation. FFSs were shown to have a positive impact on production and income among women, low-literacy, and medium land size farmers.

Similarly to DID, other studies have used panel data in which all the variables of interest are observed in several points in time. This allows accounting for individual heterogeneity and for unobservable variables such as cultural factors or, for instance, differences in agricultural practices across villages. Effects are usually estimated through individual fixed effects. Moreover, it is possible to evaluate the short and long term effects of a program. As in the simple DID model, the identification of impacts relies on the assumption of time-invariant unobservable factors. Several papers using panel data find that extension programs have a positive impact in the short run and a negligible impact in the long run. Owens et al. (2003) find that, after controlling for innate productivity characteristics and farmers' abil-

[4] In this context, project activities were not randomly allocated, therefore endogenous placement is likely to have occurred.

ity, access to agricultural extension services raises the value of crop production by about 15%. A positive impact of extension services on productivity is also found in Romani (2003) in the Ivory Coast. Maffioli et al. (2011) evaluate the effectiveness of the Farm Modernization and Development Program in Uruguay: relying on both panel data and propensity score matching, they find a positive effect of the program on improved production techniques, such as plantation density and the rate of adoption of certified varieties. Maffioli et al. (2013) find evidence that extension programs increased plantation density. However, they find no evidence of the program's impact on yields for the period under study.

Ashraf et al. (2009) is the only work using RCT reporting on an extension project implemented in Kenya and attempting to support farmers in adopting market export crops by relieving financial and informational constraints. Farmers participating in self-help groups were randomly allocated to receive complete, partial or no project services, the latter being the control group. Results suggest that after one year the program led to an increase in the production of export-oriented crops and lower marketing costs; this translated into household income gains for new adopters. However, the services collapsed one year later because farmers could not meet new EU production requirements.

4. A CASE STUDY: AGRICULTURAL EXTENSION SERVICES IN ETHIOPIA

In Bonan, Rotondi and Pareglio (2015) we contribute to the literature on the effect of extension services by evaluating a program implemented by CIAI, an Italian NGO, in a rural village in Ethiopia between 2013 and 2014. The program aimed at introducing the cultivation of horticultural gardens along with some more innovative techniques, products and inputs, and at contributing to welfare improvement through diet diversification and revenues from the sales of products. The farmers were made to attend two training sessions in 2013 and 2014, lasting three days and focused on the horticultural cultivation of vegetables and fruits. Theoretical sessions on the advantages of vegetable and fruit production and consumption to improve household welfare and diet were accompanied by the preparation of show-gardens. Farmers were also trained to the correct use of pesticides, both natural and chemical, to control for major pests and diseases. Training on soil fertility management, compost making and post-harvest practices was also provided. Participants had the possibility to receive vegetable and fruit seeds and some agricultural tools (hoes and racks). Farmers were then visited by project staff (in collaboration with the local development office) for a technical follow-up on garden maintenance and on the implementation of the techniques explained during the training sessions. During the visit the project staff made a qualitative assessment of the state of the garden and of the implementation of soil fertility management.

To assess the impact of the project, we used a mixed impact evaluation design combining across-villages comparisons, through difference-in-differences (DID) estimations, with a within village randomized control trial design. To this aim, we

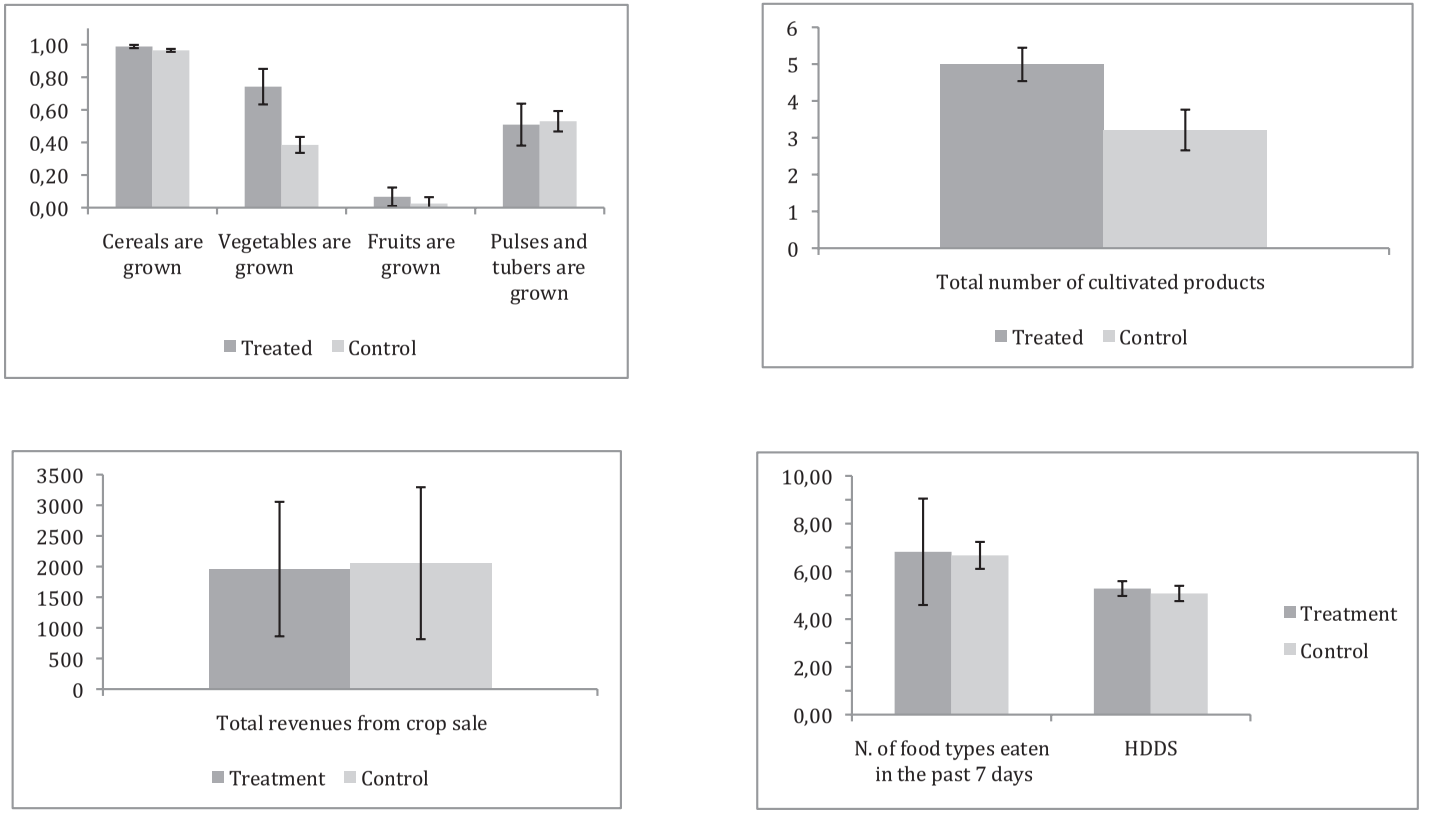
used micro-data collected through field surveys carried out in two time periods (2013 and 2014). The sampling procedure consisted of three steps: the random selection, from a list of 545 potential beneficiaries, of 250 households that actually received the treatment in the treated village; the random selection, from the pool of remaining households, of a sample of 121 households that did not receive the treatment; and, finally, a random selection, in another village with similar characteristics as the treated village, of a pure control group composed of 250 households from a list 425 farmers provided by the local administration authorities.

Empirically, we first exploited a difference-in-difference design to calculate the effect of the treatment on the outcomes of interests by comparing the average change over time in the outcomes of the treatment group with the average change over time of the pure control groups. As mentioned above, this estimator relies on the assumption that, conditional on observed characteristics, the evolution of the outcomes in treated and pure control areas would have been the same in the absence of the project. We then exploited a within-village randomization design and implemented an instrumental variable approach using as instrument the random assignment to treatment and control group. Randomization ensured that the two groups were homogeneous and ''identical on average'' on all aspects but treatment exposure, and allowed us to attribute causal effects of the project to the observed differences in outcomes. The within-village randomization approach, however, bears the risk of violation of the assumption according to which the treatment status of any unit does not affect the potential outcomes of the other units (non-interference). In fact, the control group situated in the same village of the treated group might have directly or indirectly benefited to some extent from treatment exposure, such as contacts and relationships between project beneficiaries and non-beneficiaries, which might have led to imitation processes. Assuming that such spillover effects, if any, are acting in the direction of the project's aims (e.g. positive effects), the evaluation exercise tends to underestimate the effects of the programs. In other words, this exercise gives us a lower-bound estimation of the real effect of the program on the outcomes of interest.

The main results of the analysis are summarized in Figure 1 where the final levels of treatment and control groups, together with confidence intervals, are shown[5]. Impacts are significant if differences between groups exist and confidence intervals do not overlap. We find that the project contributes to the diversification of production, with the average introduction of two more products, which had not been previously cultivated. Evidence suggests that this result is driven by an increase of approximately 30% in the number of households that grow vegetables. The results seem to show that extension services are an effective strategy to introduce vegetables in home gardens. The impact of vegetable sales from small land extensions, as proposed by the project, remains secondary and does not seem to

[5] Results reported are average of findings across different models and specifications detailed in Bonan et al. (2015).

FIGURE 1. – *Project's Impact on Agricultural Production, revenues from sale and diet diversification*



Source: Authors' elaboration from Bonan et al. (2015).

impact total sales revenues. We do not find significant changes in diet diversification levels in terms of number of different food items consumed and of Household Dietary Diversity Score (HDDS) – an index developed by the Food and Nutrition Technical Assistance Project (FANTA). However, we cannot conclude that the program failed to improve the households' nutritional status. The perishable nature of the products introduced and their seasonality, combined with the relatively small size of land dedicated to home gardens, the absence of irrigation and adequate storing technologies, are important barriers to multiple cultivation cycles and to the consequent persistence of effects on diet diversification across time. Arguably, better results would have been reached by explaining the basics of storage and irrigation technologies in the training sessions.

5. CONCLUSION

Agriculture is central in the debate on the new post-2015 development agenda, food security and poverty eradication being directly linked to agricultural production. However, evidence of the impact of development programs aimed at introducing extension services to improve agricultural productivity (DFID, 2014) is still scarce, despite their relevance and the amount of money invested.

Although the use of impact evaluation studies for the analysis of policies, programs and projects by national and international institutions, agencies, donors and NGOs, has considerably increased in the last decade, impact evaluation is still used only in a minority of cases. Adopting proper impact evaluation designs has several advantages. First of all, impact evaluation helps policy makers in understanding what works and what does not. It therefore feeds the decision-making process in terms of the efficiency and cost-effectiveness of solutions. Moreover, the design of an impact evaluation exercise helps to better clarify program objectives, in particular because it establishes well-defined measures of a program's success. It therefore forces policy makers to be accountable and to set clearer goals to ensure that the results will be highly policy relevant. Last, impact evaluation studies contribute to increasing the global knowledge of development processes. The full support of policy makers is a prerequisite for carrying out successful evaluations. Impact evaluation methods have also some relevant limitations; one over all is their inability to fit all the contexts in the same way. For example, experimental methods are particularly suitable for evaluating simple replicable interventions, with clear objectives and easy implementation protocols. They are very useful in assessing pilot projects before scaling-up. Conversely, experimental methods cannot capture general equilibrium effects. They are therefore unable to assess the consequences of projects beyond the treatment and control groups. Nevertheless, we think that considerations on cost-effectiveness and efficiency of policies and programs based on the evidence collected through impact evaluation studies are extremely important when resources are scarce as they allow identifying the best way to spend money for the goals of interest.

REFERENCES

Aker J.C. (2011). Dial ''A'' for agriculture: a review of information and communication technologies for agricultural extension in developing countries. *Agricultural Economics*, **42**(6), 631-647

Ashraf N., Giné X., Karlan D. (2009). Finding Missing Markets (and a Disturbing Epilogue): Evidence from an Export Crop Adoption and Marketing Intervention in Kenya. *American Journal of Agricultural Economics*, **91**(4), 973-990

Banerjee A., Duflo E. (2012). *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. PublicAffairs, New York, NY

Birner R., Davis K., Pender J., Nkonya E., Anandajayasekeram P., Ekboir J., Cohen M. (2009). From best practice to best fit: a framework for designing and analyzing pluralistic agricultural advisory services worldwide. *Journal of agricultural education and extension*, **15**(4), 341-355

Bonan J., Rotondi V., Pareglio S. (2015). *Impact of Home Gardening on Production and Diet Diversification. Evidence from a Field in Ethiopia*. mimeo

Christiaensen L., Demery L. (2007). Down to Earth. World Bank

Davis K., Nkonya E., Kato E., Mekonnen D.A., Odendo M., Miiro R., Nkuba J. (2012). Impact of farmer field schools on agricultural productivity and poverty in East Africa. *World Development*, **40**(2), 402-413

Dercon S., Gilligan D.O., Hoddinott J., Woldehanna T. (2009). The Impact of Agricultural Extension and Roads on Poverty and Consumption Growth in Fifteen Ethiopian Villages. *American Journal of Agricultural Economics*, **91**(4), 1007-1021

DFID (2014). Can agriculture interventions promote nutrition? Agriculture and nutrition evidence paper. DFID, London

FAO, WFP (2014). State of food insecurity in the world 2014: strengthening the enabling environment for food security and nutrition. Food and Agriculture Organization of the United Nations, Rome

Godtland E.M., Sadoulet E., Janvry A.D., Murgai R., Ortiz O. (2004). The impact of farmer field schools on knowledge and productivity: A study of potato farmers in the Peruvian Andes. *Economic Development and Cultural Change*, **53**, 63-92

Imbens G.W., Wooldridge J.M. (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, American Economic Association, **47**(1), 5-86

Krishnan P., Patnam M. (2014). Neighbors and extension agents in Ethiopia: Who matters more for technology adoption? *American Journal of Agricultural Economics*, **96**(1), 308-327

Maffioli A., Ubfal D., Baré G.V., Cerdán-Infantes P. (2011). Extension services, product quality and yields: the case of grapes in Argentina. *Agricultural Economics*, **42**, 727-734

Maffioli A., Ubfal D., Vazquez-Bare G., Cerdan-Infantes P. (2013). Improving technology adoption in agriculture through extension services: evidence from Uruguay. *Journal of Development Effectiveness*, **5**(1), 64-81

Owens T., Hoddinott J., Kinsey B. (2003). The impact of agricultural extension on farm production in resettlement areas of zimbabwe. *Economic Development and Cultural Change*, **51**(2)

Rivera W.M., Kalim Qamar M., Van Crowder L. (2001). Agricultural and Rural Extension Worldwide: Options for Institutional Reform in Developing Countries. Food and Agriculture Organization, Rome

Romani M. (2003). The impact of extension services in times of crisis: Cote d'Ivoire (1997-2000). Development and Comp Systems 0409053, EconWPA

Udry C. (2010). The economics of agriculture in Africa: Notes toward a research program. *African Journal of Agricultural and Resource Economics*, **5**(1)

van der Ban A.W., Hawkins H.S. (1996). *Agricultural Extension*. Blackwell Science, Oxford, 2nd ed

World Bank (2015). Enabling the Business of Agriculture. Progress Report. The World Bank, Washington DC