Developing a Methodology for Assessing the Impact of Farmer Field Schools in East Africa

Kristin Davis
International Food Policy Research Institute
PO Box 5689
Addis Ababa, Ethiopia
Tel. +251 11 645 8812; Fax +251 11 646 2927
k.davis@cgiar.org

Ephraim Nkonya
International Food Policy Research Institute

Abstract

Although farmer field schools (FFS) have been used worldwide in extension education for almost 20 years, most studies of their impact have been limited as to scope or rigor. This paper presents a methodology for better understanding how FFS have evolved in three countries in East Africa in response to various factors, and to see what the impact of FFS has been on poverty, innovation, empowerment, gender, productivity, and sustainability of agriculture. This methodology includes both qualitative and quantitative data collection and analysis, and incorporates propensity score matching and double difference approaches to improve the design to show impact of the FFS.

Key words: Farmer field schools, impact, Kenya

Introduction

Despite the tremendous need for extension and education approaches that contribute to poverty reduction in rural sub-Saharan Africa, many models have been tried with only limited success. One highly successful extension approach is the farmer field schools (FFS) approach. Started in Indonesia in 1989, the approach has expanded throughout many parts of sub-Saharan Africa. In Kenya alone, there are over 1,500 FFS with 34,000 farmer graduates (Abate and Duveskog, 2003).

As FFS implementation is being scaled up in Africa there are growing interest and concerns among stakeholders and donors regarding applicability, targeting, cost-effectiveness, and impact of the FFS approach. However, there are a limited number of studies documenting FFS impact in a systematic manner and are both rigorous and broad in scope (van den Berg, 2004). Because of the limited number of studies, especially in Africa, there is still much unknown about the approach and the issues pertinent to extension, such as effectiveness, sustainability, participation, and financing.

Part of the reason for the lack of empirical studies is the difficulty in tracing impact (Anderson & Feder, 2007; Purcell & Anderson, 1997). Many infrastructural variables and other factors affect agricultural performance in complex and contradictory ways, and benefits are difficult to quantify (Anderson, 2007; Birkhaeuser, Evenson, & Feder, 1991) and to attribute to programs or project interventions. Measurement challenges of several types contribute to the difficulty, and questions of representativeness abound in any attempt at grouping. Other issues that may confound studies include endogeneity in program placement and extension-farmer interactions, and selection bias. Lack of baseline data is another problem that leads to the inability to conduct FFS impact assessment.

To date, no study in East Africa has been both rigorous in terms of methods used and in terms of the scope and scale of the study. This study will help to fill this gap and to answer some of the questions plaguing FFS stakeholders.

Purpose and Objectives

In order to provide robust evidence for policy makers, donors, farmers, and implementation actors on if and how FFS can contribute to poverty alleviation, productivity, and local empowerment, the International Food Policy Research Institute plans to evaluate the East African Sub-regional Project for Farmer Field Schools, which began in 1999 in Kenya, Tanzania, and Uganda. Specific objectives of the study are found below.

(1) Determine the effectiveness of FFS in achieving outcomes regarding poverty, innovation, empowerment, gender, productivity, and/or sustainability of agriculture.
(2) Ascertain the impact that broader context (rural services, policies, markets, agricultural potential, population density) have had on FFS.
(3) Determine the impacts of household capital endowment level social characteristics on FFS (including human capital (level of education, family labor, etc), physical capital (houses, productive equipment, land, etc), social capital, and financial capital).
(4) Determine whether the poor, women and other marginalized groups participate equally in FFS.
(5) Determine how FFS have evolved within and between countries and institutional arrangements; and in particular, establish whether any learning mechanisms have been created and what factors have influenced the development of FFS.

The purpose of this paper is to explain the specific methods that will be used to address some of the impact assessment issues mentioned above. Using these specific methods will help to answer the overall research objectives above.

Methods and Data Sources

Both qualitative and quantitative methods will be used to collect and analyze data, including using document analysis, semi-structured interviews with key informants, and primary and secondary survey data. To achieve the first, third, and fourth objectives of this study, we will use mainly quantitative methods.

Impact studies of programs face three interrelated challenges – establishing a viable counterfactual (the predicted outcome in the absence of the intervention – i.e., what would have happened to the beneficiaries had they not participated in the FFS), attributing the impact to an intervention, and coping with long and unpredictable lag times (Alston and Pardey, 2001; Salter and Martin, 2001). If we measure the impact of FFS using income as an indicator of poverty, the average impact of a FFS on the beneficiaries (referred to in the impact assessment literature as the average treatment effect on the treated (ATT) is defined as the difference between the expected income earned by FFS members while participating in the FFS and the expected income they would have received if they had not participated in the FFS:

$$ATT = E(Y_1|p = 1) – E(Y_0|p = 1) \text{ …………(1)}$$

Where ATT = average impact of treatment on the treated; p = participation in the FFS (p = 1 if participated in the FFS and p = 0 if did not participate in the FFS); Y_1 = outcome (household income in this example) of the FFS beneficiary after participation in FFS; Y_0 = outcome (income) of the same beneficiary if he/she had not participated in the FFS.

Unfortunately, we cannot observe the counterfactual income of the FFS beneficiaries had they not participated in the FFS ($E(Y_0|p=1)$. Simply comparing incomes of households that are participating in the FFS and those that are not can result in serious biases, since these two groups may be quite different and hence are likely to have different incomes regardless of their participation in the FFS. Experimental and quasi-experimental methods have been used to find the groups who do not participate in the program but have socio-economic and biophysical characteristics that are similar to the participants. However, if unobservable characteristics also affect the outcomes being compared, estimates from these comparisons may be biased since the choice of comparable groups is based on observable characteristics only. The bias can be expressed by adding and subtracting $E(Y_0|p=0)$ on the right hand side of equation (1):

$$ATT = [E(Y_1|p = 1) – (E(Y_0|p = 0)] – [E(Y_0|p = 1) – (E(Y_0|p = 0)] \text{ ……..(2)}$$

The first expression (in the first square bracket) is observable since it is the difference of income of the beneficiaries and non-beneficiaries. The second expression (which is unobservable because $E(Y_0|p = 1)$ is unobservable) represents the bias resulting from estimating ATT as the first expression. This bias results because the income that non-beneficiaries receive without the program may not be equal to the income that beneficiaries would have received without the program (i.e., $E(Y_0|p = 1)$ is not equal to $(E(Y_0|p = 0))$.

Two common sources of bias are program placement or targeting bias, in which the location or target population of the program is not random, and self-selection bias, in which households choose whether or not to participate, and thus may be different in their experiences, endowments and abilities.

The most accepted method to address these biases is to use an experimental approach to construct an estimate of the counterfactual situation by randomly assigning households to treatment (beneficiary) and control (non-beneficiary) groups. Random assignment assures that both groups are statistically similar (i.e., drawn from the same distribution) in both observable and unobservable characteristics, thus avoiding program placement and self selection biases. Such an approach is not feasible in demand-driven programs in which participants make their own decisions of whether to participate and on the kind of activities to do in the learning process. Likewise, random assignment also conflicts with the nature of community-driven development programs like FFS.

Various quasi-experimental and non-experimental methods have been used to address the bias problem (for details see Rosenbaum and Rubin, 1983; Heckman, Ichimura, and Todd, 1997; 1998; and Smith and Todd, 2005). One of the most commonly used quasi-experimental methods used is the propensity score matching (PSM), which selects beneficiaries and non-beneficiaries who are as similar as possible in terms of observable characteristics expected to affect program participation as well as outcomes. This method is referred to as a “quasi-experimental” method because it seeks to mimic the approach of experiments in identifying similar “treatment” and “control” groups. However, since the comparison groups identified in PSM are not selected by random assignment, they may differ in unobserved characteristics, even though they are matched in terms of observable characteristics. The difference in outcomes between the two matched groups can be interpreted as the impact of the FFS on the beneficiaries (Smith and Todd, 2005). We used this method to estimate the ATT for impacts of the FFS on household income, productivity and other quantifiable factors that determine effectiveness of FFS.

The PSM method matches the FFS beneficiaries with comparable non-beneficiaries using a propensity score, which is the estimated probability of being included in the FFS. Only beneficiaries and non-beneficiaries that have comparable propensity scores are used to estimate the ATT. To illustrate the importance of using PSM, we compared results obtained using matched and unmatched beneficiaries and the non-beneficiaries. Matched groups are the treated (beneficiary) group that has comparable observable characteristics as the control (non-beneficiary) group. We used data from a Nigerian project (Fadama II) that offered demand-driven agricultural advisory services (See Nkonya, Phillip, Mogues, Yahaya, Adebowale, Pender, et al., 2008 for details). Table 1 shows generally similar results of matched and unmatched samples. However, some results lead to different conclusions, which highlight the bias introduced by including observations that are not comparable. For example, results from the matched samples showed no significant difference (at 10%) between the demand for agricultural marketing information by the project beneficiaries and non-beneficiaries. However, when all observations – including those which do not match – were used, project non-beneficiaries showed a significantly higher (at $p = 0.05$) demand for agricultural marketing information. This demonstrates the bias due to inclusion of observations that do not have comparable observable characteristics and underscores the need to use PSM to isolate such observations.

Table 1

Adoption and Demand for Production, Post-harvest, Financial Management and Marketing Technologies in Nigeria

Technology	Asked for the technology (matched sample)			Asked for technologies (entire sample)		
	Proportion reporting yes		*P*	Proportion reporting yes		*p*
	Fadama II beneficiaries	All non Fadama II		Fadama II beneficiaries	All non Fadama II	
Crop improved varieties	0.056	0.093	0.170	0.050	0.065	0.339
	(0.018)	(0.017)		(0.012)	(0.009)	
Soil fertility management	0.038	0.105	0.041**	0.032	0.066	0.092*
	(0.022)	(0.027)		(0.016)	(0.016)	
Livestock Management	0.047	0.041	0.795	0.046	0.037	0.595
	(0.019)	(0.015)		(0.012)	(0.010)	
Post harvest handling	0.271	0.160	0.098*	0.371	0.176	0.000***
	(0.054)	(0.041)		(0.041)	(0.029)	
Financial management	0.013	0.069	0.022**	0.024	0.059	0.037**
	(0.009)	(0.030)		(0.009)	(0.022)	
Agricultural marketing	0.016	0.059	0.199	0.030	0.082	0.031**
	(0.016)	(0.026)		(0.015)	(0.021)	

Note. Figures in parenthesis are standard errors.

PSM has some advantages over econometric regression methods since it compares only comparable observations and does not rely on parametric assumptions to identify the impacts of FFSs (Heckman, et al., 1998). However, PSM does not address selection bias due to unobservable characteristics. The beneficiary and comparison groups may differ in unobservable characteristics, even though they are matched in terms of observable characteristics (Ibid.). Econometric regression methods have been devised to address this problem, although these suffer from the problems noted above. It has been reported that the bias resulting from comparing non-comparable observations can be much larger than the bias resulting from selection on unobservables (Ibid.), although one cannot say whether this conclusion holds in general.

In this study, we addressed the problem of selection on unobservables by combining PSM with the use of the double-difference (DD) estimator. Double Difference is also referred to as Difference-in-Difference (DID) (Duflo et al., 2004). The DD estimator compares changes in outcome measures (i.e., change from before to after the FFS) between program participants and non-participants, rather than simply comparing outcome levels at one point in time. The advantage of this estimator is that it nets out the effects of any factors (whether observable or unobservable) that have fixed (time-invariant) and additive impacts on the outcome indicator (Ravallion, 2005). Thus, for example, if program participants and non-participants are different in their asset endowments (mostly observable) or in their abilities (mostly unobservable), and if these differences have an additive and fixed effect on outcomes during the time period studied, such differences will have no confounding effect on the estimated ATT. In principle, this approach can be used to assess program impacts without using PSM, and will produce unbiased estimates of impact as long as these assumptions hold. However, if the program has differential impacts on people having different wealth or other observable characteristics, the simple DD estimator will produce biased estimates if participant and non-participant households differ in these characteristics (Ibid.). By combining PSM with the DD estimator, differences in pre-FFS observable characteristics can be controlled for. There still could be a bias due to heterogeneous or time varying impacts of the unobservable differences between participants and non-participants. Such shortcomings are unfortunately inherent in all non-experimental methods of impact assessment. There is no perfect solution to these potential problems, and we believe that the PSM and DD methods will address these issues.

We used econometric analysis to achieve the first, fourth, and third objectives of this study. The major objective of the econometric analysis will be to determine the impact of village and national level factors (market access, policy, population density, etc) and household level endowment of capital on the effectiveness of FFS. Results from this analysis will help in recommending the actions required to increase the effectiveness of FFS in Sub-Saharan Africa. The econometric analysis will be based on matched (comparable) samples drawn from the PSM approach. This ensures that the results do not include groups that are not comparable. As is the case for the PSM/DD approach discussed above, this reduces the bias due to inclusion of non-comparable observations (Eliasson, 2006).

To answer objectives 2 and 5, qualitative methods will be used, including analysis of secondary data, key informant interviews, and document analysis. Project documents and other relevant literature will be examined, together with key informant interviews, to see how various policies and other factors have affected FFS. Secondary data at both the household and community level will be analyzed to see how markets, agricultural potential, population densities, and rural services affect FFS. Finally, document analysis and interviews with key informants will be used to analyzed how FFS have evolved within different institutional and other settings, and to see what learning mechanisms have been created.

Conclusions

This paper has presented a methodology for better understanding how FFS have evolved in East Africa in response to various factors, and to see what the impact of FFS has been on poverty, innovation, empowerment, gender, productivity, and sustainability of agriculture. This methodology includes both qualitative and quantitative data collection and analysis, and incorporates propensity score matching and double difference approaches to improve the design.

Educational Importance and Implications

Understanding the impact of education and extension programs is crucial. Program managers and extension personnel, donors, governments, and other stakeholders would all like to know if and how well programs are working. This paper demonstrates a methodology for showing the impact of extension and education programs in developing countries.

References

Abate, A. & Duveskog, D. (2003). *Application of the farmer field school approach in Kenya. In: FAO-KARI-ILRI. Farmer field schools: The Kenyan Experience*. Report of the Farmer Field School Stakeholders' Forum, 27 March, International Livestock Research Institute, Nairobi, Kenya.

Alston, J.M. & Pardey, P.G. (2001). Attribution and other problems in assessing the returns to agricultural R&D. *Agricultural Economics*. 25: 141-152.

Anderson, J. R. & Feder, G. (2007). Agricultural extension. In: Evenson, R. and P. Pingali (Eds.). *Handbook of agricultural economics Vol. 3.* Elsevier B.V.

Birkhaeuser, D., Evenson, R. E., & Feder, G. (1991). The economic impact of agricultural extension: A Review. *Economic Development and Cultural Change 39*(3): 607-640

Duflo, E., Mullainathan, S. & Bertrand, M. (2004). How much should we trust difference in difference estimates? *Quarterly Journal of Economics*, 119(1): 249-275.

Eliasson, K., (2006). *How robust is the evidence on the returns to college choice? Results using Swedish administrative data*. Working Paper. Department of Economics, Umeå University, and National Institute for Working Life.

Heckman, J, Ichimura, H., Smith, J, & Todd, P. (1998). Characterizing selection bias using experimental data. *Econometrica*. 66: 1017-99.

Heckman, J., Ichimura, H., & Todd, P. (1997). Matching as an econometric evaluation estimator. *Review of Economic Studies* 65(2): 261–294.

Nkonya, E., Phillip, D. Mogues, T. Yahaya, M. Adebowale, G. Pender, J. et al. (in press). From the ground up: Targeting the poor through a community driven development project in Nigeria. Discussion Paper, International Food Policy Research Institute.

Purcell, D. L. & Anderson, J. R. (1997). *Agricultural extension and research: Achievements and problems in national systems*. Washington, DC: The World Bank.

Ravallion, M. (2005). *Evaluating anti-poverty programs*. World Bank Policy Research Working Paper 3625. The World Bank, Washington, D.C.

Rosenbaum, P. R. & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1): 41–55.

Salter, A.J. and B.R. Martin. (2001). The economic benefits of publicly funded basic research: a critical review. *Research Policy*. 30: 509–532.

Smith, J. & Todd, P. (2001). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics*, 125(1-2): 305-353.

van den Berg, Henk. (2004). *IPM FFS: A synthesis of 25 impact evaluations*. Report prepared for the Global IPM Facility. Wageningen: The Netherlands.