Impact Assessment of Livestock Research and Development in West Africa: A Propensity Score Matching Approach

Sabine Liebenehm
Leibniz Universität Hannover, Germany

Hippolyte Affognon
International Livestock Research Institute, Nairobi, Kenya

Hermann Waibel
Leibniz Universität Hannover, Germany

Abstract

International agricultural research often uses quasi-experimental designs when implementing on-farm research and development activities. Therefore, impact assessment methodologies are needed, which are effective in circumventing the resulting selection bias inherent in such research designs. This paper applies propensity score matching (PSM) as one way to measure the impact of ILRI-led research activities to control African animal trypanosomosis (AAT) in West Africa. Data were collected from 508 farmers in Mali and Burkina Faso. Results indicate significant improvements in farmers' knowledge. The paper adds to the methodology of PSM in impact assessment by emphasizing on the quality of different matching algorithms and on the sensitivity of impact estimates.

Keywords: livestock disease, knowledge, propensity score matching
JEL: Q12, D83, C14

1 Introduction

One of the major research and development (R&D) activities of the International Livestock Research Institute (ILRI) in the cotton zone of West Africa is to achieve effective control of African Animal Trypanosomosis (AAT). AAT is the most important cattle disease with major economic impacts in terms of output losses and it is a major limiting factor for agricultural development in West Africa (SWALLOW, 1999). Treating cattle with trypanocidal drugs is the most common control strategy. Because of frequently inappropriate use, pathogens can develop resistance to the drugs (GEERTS and HOLMES, 1998). Through limited market size in West Africa, it is unlikely that any new trypanocide will be developed in the near future. To meet this challenge ILRI promoted the concept of rational drug use (RDU) as a strategy to maintain the susceptibility of pathogens to trypanocides (AFFOGNON, 2007). RDU is one possibility to prolong the effectiveness of existing drugs and to increase productivity through a reduction of output loss. ILRI in cooperation with National Agronomic Research Systems (NARS), extension services and German universities has undertaken measures to outreach the RDU concept by providing farmers with adequate information material in local language and by an advisory approach which included the diagnosis and judicious use of trypanocidal drugs jointly implemented by researchers, extension workers, 'paravets'[1] and farmers in selected villages (GRACE et al., 2008).

In ex-post impact assessment of natural resource management type of technologies, the examination of a change in knowledge of improved practices is a good starting point to measure the effectiveness of extensional activities (BIRKHAEUSER et al., 1991; ZILBERMAN and WAIBEL, 2007). Moreover, the measurement of knowledge change has the benefit of being independent from factors like prices or input supplies that may vary over time and across locations (FEDER and SLADE, 1986).

The objective of this paper is to measure the effect of ILRI-led R&D activities on farmers' disease knowledge and management practices at a time when research activities have ended, but up-scaling has not yet being started. Hence, the study serves as a baseline to appraise the current and expected effects of livestock research activities on knowledge.

A procedure is introduced that can help to overcome the problem arising from quasi-experimental designs often found in field based agricultural R&D activities. In the case of ILRI-led R&D farmers could decide by themselves to use the information provided. Hence, participating farmers may differ in various factors from non-participants. In the conventional impact assessment model, where the with and without treatment scenarios are compared, this selection bias can lead to misguided conclusions. Hence, this paper applies propensity score matching (PSM) to circumvent these limitations. The PSM method has been applied in previous impact assessment research (e.g. GODTLAND et al., 2004; KRASUAYTHONG, 2008). This paper adds to this literature by including a quality check of different matching algorithms, as well as a sensitivity analysis to control for unobservable influences.

In the next section, data collection procedures are desribed. In section three the methodology of PSM including quality check and sensitivity analysis is presented. Section four illustrates the results and in the last section conclusions are drawn.

[1] Paravets are farmers selected by the community and trained in primary animal health services.

2 Data Collection Procedures

ILRI-led R&D activities took place in villages across south-eastern Mali and south-western Burkina Faso. In order to conduct ex-post impact assessment, the same project villages in the study area had been revisited. All farm households in the respective villages that possessed cattle were selected. To distinguish between project participants and the control group, farmers were asked if they attended R&D activities led by ILRI. A knowledge, attitude and practice (KAP) questionnaire was administered with household heads, generally responsible for livestock production and animal health management (GRACE et al., 2009). This form represents an appropriate tool to evaluate qualitative issues by quantitative information (HAUSMANN-MUELA et al., 2003). Based on this KAP questionnaire, scores had been allocated according to three different knowledge categories:

i. disease specific knowledge, comprising signs, causes, possibility of animal re-infection after being cured and animals' susceptibility to the disease;
ii. curative treatment knowledge and actual control actions in case of trypanosomosis occurrence, including the quality and quantity of trypanocides for treatment; and
iii. preventive treatment knowledge and actual preventive strategies applied, also involving cattle husbandry and medical management.

Summing up all points from the three categories above gives the total knowledge score. Following the procedure applied in integrated pest management for crops, knowledge categories are calculated in percentage of the maximum possible score (GODTLAND et al., 2004).

Originally developed in French language, trained interviewers administered the questionnaire in the respective local languages, i.e. *Bambara* in Mali and *Djoula* in Burkina Faso[2]. The questionnaire included both open-ended and pre-coded questions. For visual support picture cards were used for knowledge questions. In total, data from 508 cattle farmers were included in the ex-post impact analysis, whereby 211 respondents participated and 297 did not attend the former research activities (LIEBENEHM et al., 2011).

2 Questionnaires were administered in French due to lack of expertise for local languages with limited possibilities of written expressions.

3 Methodology of Propensity Score Matching

In order to draw more precise conclusions about the impact of ILRI-led R&D activities on farmers' knowledge, it becomes necessary to circumvent the selection bias on observables. Matching on the probability of participation, given all observable treatment-independent covariates X solves the problem of selection bias. The propensity score of vector X can be defined as:

(1) $$P(X) = \Pr(Z = 1 \mid X),$$

where Z denotes the participation indicator equaling one if the individual participates, and zero otherwise. Given that the propensity score is a balancing score, the distribution of observables X will be the same for both participants and non-participants. Consequently, the differences between the groups are reduced to the attribute of treatment assignment, and unbiased impact estimates can be produced (ROSENBAUM and RUBIN, 1983) by the following four steps.[3]

Firstly, the probability of participation is predicted by a binary response model with appropriate observable characteristics. Various methods to predict propensity score produce similar impact estimates (TODD, 1995). For computational simplicity a logit model will be applied here. The propensity score can then be defined as:

(2) $$P(X) = \Pr(Z = 1 \mid X) = F(\beta_1 x_1 + ... + \beta_i x_i) = F(X\beta) = e^{X\beta},$$

where F(·) produces response probabilities strictly between zero and one.

The performance difference between treatment and control groups is estimated by the average treatment effect on the treated (ATT) in a second step. The true ATT, based on PSM, can be written as:

(3) $$ATT_{PSM} = E_{P(x)}\{E(Y_1 \mid Z = 1, P(X)) - E(Y_0 \mid Z = 0, P(X))\},$$

where $E_{P(X)}$ represents the expectation with respect to the distribution of propensity score in the entire population. The true ATT indicates the mean difference in knowledge test scores between participants and non-participants, who are identical in observable characteristics and adequately weighted by a balanced probability of participation. An adequate match of a participant with his counterfactual is achieved, as long as they are identical in their observable characteristics. In order to obtain such matched pairs three different matching methods that vary in terms of bias and efficiency are applied (CALIENDO and KOPEINIG, 2005). Firstly, nearest neighbor

[3] For more details on underlying assumptions in the PSM methodology see LIEBENEHM et al. (2009).

matching (NNM) involves the selection of one non-participant with the propensity score closest to that of the respective participant. NNM will cause no concern as long as the distribution of propensity scores of the pair is similar (SMITH and TODD, 2005). Secondly, radius matching (RM) involves all neighbors within a maximum propensity score distance (caliper), a priori defined. Here, poor matches through too distant neighbors are avoided (DEHEJIA and WAHBA, 2002; SMITH and TODD, 2005). Thirdly, kernel-based matching (KM), a non-parametric matching estimator, includes all individuals of the underlying sample of non-participants and weights more distant observed characteristics among both groups down (HECKMAN et al., 1997; 1998). Hence, kernel-based matching on all control units indicates a lower variance (CALIENDO and KOPEINIG, 2005).

The third step is to check the matching estimators' quality by standardized differences in observables' means between participants and non-participants. The standardized difference in percent after matching represents, for a given independent covariate X, the difference in sample means in the participating ($\overline{X}_1$) and matched non-participating ($\overline{X}_0$) sub-samples as a percentage of the square root of the average sample variances (s_1^2 and s_0^2) (ROSENBAUM and RUBIN, 1985):

$$(4) \qquad SD = \left| 100 * \frac{(\overline{X}_1 - \overline{X}_0)}{(0.5 * (s_1^2 + s_0^2))^{\frac{1}{2}}} \right| .$$

Although there exists no clear threshold of successful or failed matching, a remaining bias below 5% after matching is accepted as an indication that the balance among the different observable characteristics between the matched groups is sufficient (DIPRETE and GANGL, 2004; CALIENDO and KOPEINIG, 2005).

Finally, in consideration of the quasi-experimental design of the ILRI-led R&D activity, it might be possible that unobservable factors like farmers' intrinsic motivation and specific abilities or preferences, had affected the participation decision. This problem of hidden bias is circumvented by the bounding approach following ROSENBAUM (2002). The logit model to estimate propensity score (equation 2) is complemented by a vector U containing all unobservable variables and their effects on the probability of participation captured by γ:

$$(5) \qquad P(X) = \Pr(Z = 1 \mid X) = F(X\beta + U\gamma) = e^{X\beta + U\gamma} .$$

Sensitivity analysis examines how strong the influence of γ on the participation process needs to be, in order to attenuate the impact of participation on potential outcomes (ROSENBAUM, 2002). For the sake of simplicity, it is assumed that the unobservable variable is a binary variable taking values zero or one (AAKVIK, 2001). The following bounds on the odds ratio of the participation probability of both individuals are applied:

$$\text{(6)} \qquad \frac{1}{e^{\gamma}} \leq \frac{P(X_m)(1-P(X_n))}{P(X_n)(1-P(X_m))} \leq e^{\gamma}.$$

Both individuals have the same probability of participation, provided that they are identical in X, only if $e^{\gamma}=1$ (ROSENBAUM, 2002). If e^{γ} is close to one and changes the inference about the treatment effect, the impact of participation on potential outcomes is said to be sensitive to hidden bias. In contrast, insensitive treatment effects would be obtained if a large value of e^{γ} does not alter the inference about treatment effects (AAKVIK, 2001). In this sense, e^{γ} can be interpreted as a measure of the degree of departure from a study that is free of unobservable selection bias (ROSENBAUM, 2002).

Following these methodological steps, the next section presents the results.

4 Results

Table 1 summarizes the average differences in knowledge scores and independent observable characteristics between participants and non-participants. Overall, the difference in means shows that the level of knowledge of all cattle farmers in the sample is very low with average test scores ranking from 13% to 25% of maximum score. However, participating farmers reach significantly higher knowledge scores in all categories than those, who had not participated. On average the differences range from 2.6% in the category of knowledge on the disease to 5.3% in the category of disease control. The total knowledge score shows a difference in means of about 3.7% between the participant and the non-participant group.[4] The examination of selected observable characteristics shows that there are significant differences in means of household size and number of children going to school. Furthermore, participating household heads are on average almost three years older than their counterparts. According to farm characteristics, participants have significantly more cattle and own more means of transport (like motorbikes, bicycles or carts) than non-participating households. With regard to epidemiological factors, the proportion of participants reporting AAT is significantly higher by 8%, but there is no significant difference regarding the perception of drug resistance between the two groups, as indicated by the realization of treatment failures after administration of trypanocides (isometamidium

[4] More detailed descriptive analysis is provided by LIEBENEHM et al. (2011).

or diminazene). Hence, observable participation incentives can be identified, which underlines the possibility that selective placement exists and therefore the need to apply propensity score matching.

Table 1. Comparison of mean knowledge score and observable characteristics across participants and non-participants (N = 508)

	Participants	Non-participants	Difference in means
Knowledge scores in % of maximum scores			
Knowledge score on disease	25.30	22.68	2.62***
Knowledge score on control	23.54	18.21	5.33***
Knowledge score on prevention	16.01	12.68	3.33***
Total knowledge score	20.81	17.14	3.67***
Household characteristics			
Household size	19.04	16.25	2.79***
Dependency ratio	0.41	0.46	-0.05
Number of children at school	4.54	3.1	1.43***
Age of household head	55.77	53.15	2.62*
Formal education of household head in years	0.75	1.07	-0.32
Farm characteristics			
Number of cattle	14.1	9.84	4.26**
Mixed farming experience of household head in years	2.05	1.47	0.58
Number of means of transport	6.67	5.56	1.11**
Epidemiological factors			
Perception of disease in %	75.36	67.34	8.02*
Perception of drug resistance in %	82.94	86.2	-3.26
Observations	**211**	**297**	

Note: Data are compared using two-tailed *t* test with $^*p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$.
Source: own survey

In accordance with chosen characteristics that capture relevant observable differences between participants and non-participants the probability of participation is predicted. Table 2 reports the results from the logit model (see equation 2), while the estimated coefficients are expressed in terms of odds of *Z=1*. Examining single observables, it is shown that the dependency ratio of the household, cattle herd size, farming experience, perception of drug resistance and the country of origin in particular significantly influence the participation decision.

Table 2. Logit model to predict the probability of participation conditional on selected observables

Dependent variable: Participation (*Z=1*) **Covariates *X***		**Odds ratio**	**Marginal effects**
Household size		1.013	0.003
Dependency ratio		0.532*	-0.152*
Number of children at school		1.048	0.011
Age of household head		0.996	-0.001
Formal education of household head		0.907	-0.024
Quadratic term of education of household head		1.006	0.002
Number of cattle		1.012**	0.004**
Mixed farming experience of household head		1.843***	0.147***
Number of means of transport		1.043	0.01
Perception of disease dummy (1 = AAT)		1.256	0.054
Perception of drug resistance dummy (1 = Resistance)		2.264***	0.182***
Country dummy (1 = Burkina Faso)		0.208***	-0.371***
Observations	508		
Summary statistics			
Log-Likelihood	-295.82584		
Pearson Chi2(495)	520.28	Prob>Chi2	0.2086
Pseudo R-squared	0.142		
Area under ROC curve	0.7462		

Note: Covariates had been controlled for endogeneity; *p<0.1, **p<0.05, ***p<0.01.
Source: own survey

Based on the predicted probability of participation, the impact of the intervention on farmers' knowledge test scores is estimated by the ATT (see equation 3). Having ensured that observations are ordered randomly and that there are no large disparities in the distribution of propensity scores, nearest neighbor matching yields the highest and most significant treatment effect estimate in all four outcome categories (table 3).

Table 3. Estimated impacts of trypanocide resistance research activities on farmers' knowledge using different matching algorithms

	Knowledge score in % of maximum scores		Average treatment effect on the treated
	Participants	Non-participants	
Nearest neighbor matching	Using the single closest neighbor		
Knowledge score on disease	25.3	22.93	2.37***
Knowledge score on control	23.54	19.29	4.25***
Knowledge score on prevention	16.01	13.0	3.01***
Total knowledge score	20.81	17.65	3.16***
Observations	211	211	
Radius matching	Using all neighbors within a caliper of 0.01		
Knowledge score on disease	25.04	23.22	1.82**
Knowledge score on control	23.17	19.27	3.9***
Knowledge score on prevention	15.79	13.18	2.6***
Total knowledge score	20.54	17.81	2.73***
Observations	194	294	
Kernel-based matching	Using a biweight kernel function and a smoothing parameter of 0.06		
Knowledge score on disease	25.28	23.37	1.91**
Knowledge score on control	23.55	19.91	3.64***
Knowledge score on prevention	16.03	13.18	2.85***
Total knowledge score	20.81	18.03	2.78***
Observations	210	293	

Note: *p<0.1, **p<0.05 and ***p<0.01.
Source: own survey

The nearest neighbor estimate of the average total knowledge gain due to participation is about 3.16%. However, since this method produces relatively poor matches due to the limitation of information, attention should be focused on the other two matching algorithms. Here, the estimated impacts of participation on knowledge score are lower regarding the respective categories. Following the radius matching algorithm the difference in total knowledge scores, as a percentage of the maximum score achieved, is about 2.73%. Moreover, the estimated treatment effect in the category of curative control knowledge and action accounts for 3.9%. Kernel-based matching produces the highest impact estimate for curative treatment knowledge and actual executed control strategies. Similarly to the radius matching estimator for the total score, the kernel-based matching algorithm produces a significant average treatment effect on the treated of 2.78% at the 1% significance level.

Consequently, it can be confirmed that trypanocide resistance research activities do in fact generate significant gains in farmers' knowledge on AAT and induce farmers to improve both the curative and preventive control strategies.

Checking the imbalance of single observable characteristics in the third step (see equation 4) shows that the matching quality of radius matching and kernel-based matching is much higher than that of the simple method of choosing the only closest neighbor with respect to the propensity score (table 4).

Table 4. Imbalance test results of observable covariates for three different matching algorithms using standardized difference in percent

Covariates *X*	Standardized differences in % after		
	Nearest neighbor matching	Radius matching	Kernel-based matching
Household size	25.9	3.5	1.2
Dependency ratio	15.3	4.0	6.3
Number of children at school	47.3	9.9	13.0
Age of household head	36.7	5.8	9.5
Formal education of household head	3.2	3.5	1.0
Quadratic term of education of household head	3.4	7.0	2.7
Number of cattle	30.8	5.2	2.7
Mixed farming experience of household head	33.5	6.9	0.5
Number of means of transport	25.9	1.2	4.7
Perception of disease dummy (1 = AAT)	12.8	1.8	2.0
Perception of drug resistance dummy (1 = resistance)	22.8	4.4	1.0
Country (1 = Burkina Faso)	101.4	5.4	5.6
Mean absolute standardized difference	29.92	4.89	4.19
Median absolute standardized difference	25.91	4.83	2.7
Observations	422	488	503

Source: own survey

However, radius matching produces high biases for two variables, i.e. number of children at school and farming experience, above the reference line of 5%. Furthermore, the only variable that seems to be unbalanced between participants and non-participants after kernel-based matching, besides number of children at school, is the covariate that indicates the age of the household head. Nevertheless, the summary statistics for the overall balance of all covariates between participants and non-participants confirms the higher quality of kernel-based matching and radius matching. Both the mean and the median of the absolute standardized difference after matching are below the threshold of 5%.

Finally, an appropriate control strategy for hidden bias is to examine the sensitivity of significance levels. Here, several values of e^{γ} bounds are calculated on the significance level of impact estimates (see equation 6). The question arises at which critical impact level of the unobservable variables the inference about the treatment effect on knowledge will be undermined, as indicated by the loss of significance (DIPRETE and GANGL, 2004). Table 5 compares the sensitivity of treatment effects on different knowledge scores among the three introduced matching algorithms. Overall, robustness results produced by Rosenbaum's bounds are quite similar.

Table 5. Sensitivity analysis with ROSENBAUM'S bounds on probability values

	Upper bounds on the significance level for different values of e^{γ}				
	e^{γ}=1	e^{γ}=1.25	e^{γ}=1.5	e^{γ}=1.75	e^{γ}=2
Nearest neighbor matching	Using the single closest neighbor				
Knowledge score on disease	0.0001	0.0072	0.0871	0.327	0.6324
Knowledge score on control	<0.0001	0.0031	0.0494	0.2284	0.5151
Knowledge score on prevention	<0.0001	<0.0001	0.0018	0.0211	0.1009
Total knowledge score	<0.0001	<0.0001	0.004	0.0074	0.0465
Radius matching	Using all neighbors within a caliper of 0.01				
Knowledge score on disease	0.0005	0.0255	0.1884	0.505	0.785
Knowledge score on control	<0.0001	0.0009	0.019	0.1149	0.3267
Knowledge score on prevention	<0.0001	<0.0001	0.0015	0.0171	0.0832
Total knowledge score	<0.0001	<0.0001	0.0007	0.0099	0.0545
Kernel-based matching	Using a biweight kernel function and a smoothing parameter of 0.06				
Knowledge score on disease	0.0001	0.012	0.1254	0.4131	0.7202
Knowledge score on control	<0.0001	0.0008	0.0194	0.1241	0.3555
Knowledge score on prevention	<0.0001	<0.0001	0.0001	0.003	0.023
Total knowledge score	<0.0001	<0.0001	<0.0001	0.0017	0.0144

Source: own survey

Kernel-based matching produces the most robust treatment effect estimates with respect to hidden bias, especially for preventive knowledge and action, as well as for total knowledge score. Matched pairs might differ by up to 100% (e^{γ}=2) in unobservable characteristics, while the impact of participation on preventive treatment knowledge, as well as on total knowledge, would still be significant at a level of 5% (p-value $= 0.023$ and p-value $= 0.0144$, respectively). The same categories of knowledge score are robust to hidden bias up to an influence of e^{γ}=2 at a significance level of 10% following the radius matching approach. The nearest neighbor matching algorithm is robust to selection bias on unobservable characteristics up to an impact level of e^{γ}=2. The estimated treatment effects on knowledge about trypanosomosis, as well as on the curative knowledge and action, are sensitive to hidden bias, at a smaller unobservable impact level of e^{γ}=1.5. These sensitivity results indicate information about uncertainty in matching estimators, although it has to be considered that these results presented here are worst-case scenarios (ROSENBAUM, 2002).

5 Conclusions

Impact assessment based on simple treatment-control comparisons can be imprecise because of compounding differences in individuals. Propensity score matching can help to overcome the selection bias that arises from the quasi-experimental design often found in international agricultural research activities. Through this methodology the difference between matched participants and non-participants is solely attributed to treatment and reliable impact estimates can be obtained. Based on the quality check by standardized differences and the control of unobservables by ROSENBAUM'S bounds, significant and robust differences between matched participants and non-participants with respect to cattle farmers' knowledge are identified. Hence, it can be concluded that improvements in farmers' knowledge is attributable directly to participation in ILRI-led R&D activities. The strongest effect of the outreach activities is related to farmers' knowledge how to treat AAT, accounting for approximately 4% according to different matching algorithms. Significant knowledge advancements on preventive control strategies of around 3% are also identified. These improvements in managerial know-how and skills demonstrate the effectiveness of the R&D activities considering that the technology is highly complex. Moreover, the magnitude of farmers' knowledge advancement in livestock disease management is similar to what has been found in other knowledge-based natural resource management technologies such as pest management in vegetable production in Thailand (e.g. KRASUAYTHONG, 2008).

In conclusion, it is shown that providing cattle farmers with access to appropriate disease control information has been helpful in increasing farmers' knowledge of trypanosomosis and in improving their treatment and prevention practices. In consideration of cattle's critical role in rural households with mixed farm production, as a next step, it will be useful to analyze the linkage between practice change and the increase in productivity.

References

AAKVIK, A. (2001): Bounding a matching estimator: The case of a Norwegian training program. In: Oxford Bulletin of Economics and Statistics 63 (1): 115-143.

AFFOGNON, H. (2007): Economic analysis of trypanocide use in villages under risk of drug resistance in West Africa. PhD diss. Leibniz Universität Hannover, Germany.

BIRKHAEUSER, D., R.E. EVENSON and G. FEDER (1991): The economic impact of agricultural extension: A review. In: Economic Development and Cultural Change 39 (3): 607-650.

CALIENDO, M. and S. KOPEINIG (2005): Some practical guidance for the implementation of propensity score matching. Discussion Paper no. 1588. Institute for the Study of Labor (IZA), Bonn, Germany.

DEHEJIA, R.H. and S. WAHBA (2002): Propensity score-matching methods for non-experimental causal studies. In: The Review of Economics and Statistics 84 (1): 151-161.

DIPRETE, T.A. and M. GANGL (2004): Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. In: Sociological Methodology 34 (1): 271-310.

FEDER, G. and R. SLADE (1986): A comparative analysis of some aspects of the training and vist system of agricultural extension in India. In: Journal of Development Studies 22 (2): 407-428.

GEERTS, S. and P.H. HOLMES (1998): Drug management and parasite resistance in bovine trypanosomiasis in Africa. PAAT technical and scientific series 1. Food and Agriculture Organization (FAO), Rome, Italy.

GODTLAND, E., E. SADOULET, A. DE JANVRY, R. MURGAI and O. ORTIZ (2004): The impact of Farmer-Field-Schools on knowledge and productivity: a study of potato farmers in the Peruvian Andes. In: Economic Development and Cultural Change 53 (1): 63-92.

GRACE, D., T. RANDOLPH , O. DIALL and P.H. CLAUSEN (2008): Training farmers in rational drug-use improves their management of cattle trypanosomosis: a cluster-randomised trial in South Mali. In: Preventive Veterinary Medicine 83 (1): 83-97.

GRACE, D., T. RANDOLPH, H. AFFOGNON, D. DRAMANE, O. DIALL and P.H. CLAUSEN (2009): Characterisation and validation of farmers' knowledge and practice of cattle trypanosomosis management in the cotton zone of West Africa. In: Acta Tropica 111 (2): 137-143.

HAUSMANN-MUELA, S., J. MUELA RIBERA and I. NYAMONGO (2003): Health-seeking behavior and the health system response. DCPP Working paper no. 14. London School of Hygiene and Tropical Medicine, London, UK.

HECKMAN, J.J., H. ICHIMURA and P.E. TODD (1997): Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. In: The Review of Economic Studies 64 (4): 605-654.

– (1998): Matching as an econometric evaluation estimator. In: The Review of Economic Studies 65 (2): 261-294.

KRASUAYTHONG, T. (2008): Technologies adoption in vegetable production in Northern Thailand. PhD diss. Leibniz Universität Hannover, Germany.

LIEBENEHM, S., H. AFFOGNON and H. WAIBEL (2009): Impact assessment of agricultural research in West Africa: an application of the propensity score matching methodology. International Association of Agricultural Economists 2009 Conference, August 16-22, 2009, Beijing, China. URL: http://purl.umn.edu/50829.

– (2011): Collective livestock research for sustainable disease management in Mali and Burkina Faso. In: International Journal of Agricultural Sustainability 9 (1): 212-221.

ROSENBAUM, P.R. and D.B. RUBIN (1983): The central role of propensity score in observational studies for causal effects. In: Biometrika 70 (1): 41-55.

– (1985): Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. In: American Statistician 39 (1): 33-38.

ROSENBAUM, P.R. (2002): Observational Studies. Springer, New York.

SMITH, J. and P.E. TODD (2005): Does matching overcome LaLonde's critique of non-experimental estimators? In: Journal of Econometrics 125 (1-2): 305-353.

SWALLOW, B.M. (1999): Impacts of trypanosomiasis on African agriculture. Review paper for the Programme Against African Trypanosomiasis (PAAT). International Livestock Research Institute (ILRI), Nairobi, Kenya.

TODD, P.E. (1995): Matching and local linear regression approaches to solving the evaluation problem with semiparametric propensity score. Mimeo, University of Chicago.

ZILBERMAN, D. and H. WAIBEL (2007): Advances in impact assessment of natural resources management research. In: Quarterly Journal of International Agriculture 47 (4): 395-420.

Contact author:
Sabine Liebenehm
Leibniz Universität Hannover, Institute of Development and Agricultural Economics,
Faculty of Economics and Management, Königsworther Platz 1, 30167 Hannover, Germany
e-mail: liebenehm@ifgb.uni-hannover.de