SCHRIFTEN DER GESELLSCHAFT FÜR WIRTSCHAFTS- UND SOZIALWISSENSCHAFTEN DES LANDBAUES E.V.

Liebenehm, S., Affognon, H., Waibel, H.: Assessing the impact of agricultural research on cattle farmers' knowledge about African animal trypanosomosis: an application of the propensity score matching approach. In: Loy, J.-P., Müller, R.A.E.: Agrar- und Ernährungsmärkte nach dem Boom. Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V., Band 45, Münster-Hiltrup: Landwirtschaftsverlag (2010), S. 387-396.

Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V., Bd. 45, 2010: 387-396

ASSESSING THE IMPACT OF AGRICULTURAL RESEARCH ON CATTLE FARMERS' KNOWLEDGE ABOUT AFRICAN ANIMAL TRYPANOSOMOSIS: AN APPLICATION OF THE PROPENSITY SCORE MATCHING APPROACH

Sabine Liebenehm[1], Hippolyte Affognon[2], Hermann Waibel[1]

Abstract

Ex-post evaluation of agricultural research is aimed to empirically provide evidence of past investments' effectiveness. This paper is intended to measure the immediate impact of livestock research activities on cattle farmers' knowledge about trypanosomosis and its therapeutic and preventive control strategies. According to the quasi-experimental design of the intervention, it is shown that its impact will be adequately estimated by propensity score matching (PSM). Based on data collected according to a knowledge, attitude and practice (KAP) questionnaire in the region of Kénédougou, results indicate a significant gain in farmers' know-how due to participation in livestock research activities.

Keywords

African animal trypanosomosis (AAT); knowledge, attitude and practice (KAP); propensity score matching (PSM)

1 Introduction

In tropical Africa livestock plays a critical role as a source of income, provider of draft power, human nutrition and organic fertilizer. A severe constraint to livestock, especially to cattle, is African animal trypanosomosis (AAT). The disease lowers livestock output such as milk and meat on the short term and reduces the animal's capability. It has been estimated that the disease causes an annual production loss of up to US$ 4.5 billion (BUDD, 1999). Research by the International Livestock Research Institute (ILRI) developed technologies for integrated disease control based on the principle of rational drug use. One example is ILRI's research on trypanocide resistance "Improving the management of trypanocide resistance in the cotton zone of West Africa", in the region of Kénédougou from June 2003 to May 2004. So far little is known on the impact of these technologies on improving farmers' knowledge and capacities to achieve a better level of disease control. The aim of this study is to analyse the effect of this livestock research project on farmers' knowledge and practices change of AAT that involves better diagnosis, as well as curative and preventive control strategies. As in many natural research management projects, part of the project design has been the extension to deliver the technology to farmers (ZILBERMAN & WAIBEL, 2007). Concrete information about correct disease diagnosis and management practices was provided to cattle farmers by researchers, veterinary and para-veterinary services (AFFOGNON, 2007). The central hypothesis of this study is that the research project triggered change in farmers' behaviour, which in turn enhanced their performance in managing the disease.

As often neglected in other research programs, impact assessment has not been part of the research design (RAITZER & KELLY, 2008). Hence, the intervention with a non-randomised selection of villages and farmers follows a quasi-experimental design. Therefore, this paper uses an application of the propensity score matching (PSM) approach in order to formulate a

[1] Institute of Development and Agricultural Economics, Leibniz University Hannover, Germany
[2] International Livestock Research Institute (ILRI), Nairobi, Kenya

meaningful counterfactual and establish causality between the potential outcomes, i.e. the difference in knowledge score between treatment and control group of the research intervention (BAKER, 2000).

In the following, procedures of sample selection and data collection are provided. Thereafter, the methodology of PSM including a sensitivity analysis is described and based on its implementation results are discussed. Finally, some conclusions are drawn.

2 Survey design and sampling

In order to measure the impact of the research activity on farmers' knowledge the project villages in the region of Kénédougou, common to south-eastern Mali and south-western Burkina Faso, were revisited from October to December 2007. The household head, respectively the decision maker, who is responsible for livestock production and animal health management, was asked to take a specific knowledge test about trypanosomosis. In addition, all those farm households in the respective villages were selected, if they posses cattle, at least one animal. Originally developed in French, trained interviewers conducted the survey, in local language (that is *Bambara* in Mali and *Djoula* in Burkina Faso), and in turn filled in the questionnaire in French. Questions were applied in open-ended manner, followed by option lists and using picture cards as visual support. In total, data from 508 cattle farmer were collected.

3 Methodology

Generally, in order to infer the impact of an intervention on individual outcome, it is necessary to draw a counterfactual scenario about the outcome performance in absence of the intervention. The challenge lies in the creation of a suitable comparison group among a large group of non-participants who are as similar as possible to the participating group to obtain unbiased outcome estimates (CALIENDO & KOPEINIG, 2005).

ROSENBAUM AND RUBIN (1983) suggest therefore matching on the probability of participation, given all observable treatment-independent covariates X. The propensity score of vector X can be defined as:

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

where Z denotes the participation indicator equalling one if the individual participates, and zero otherwise. Given that the propensity score is a balancing score, the probability of participation conditional on X will be balanced such that the distribution of observables X will be the same for participants and non-participants. Consequently, the differences between both groups are reduced to the only attribute of treatment assignment and unbiased impact estimates can be produced (ROSENBAUM & RUBIN, 1983). The counterfactual group can be identified if potential outcomes Y_1 (Y_0) of participants (non-participants) are independent of participation, conditional on observables X:

(2) $$Y_0, Y_1 \perp Z \mid X, \forall X .$$

This conditional independence assumption indicates that the selection is exclusively based on the vector of observables X that determines the propensity score (ROSENBAUM & RUBIN, 1983; CALIENDO & KOPEINIG, 2005). Additionally, in order to ensure randomised selection the common support condition needs to be applied:

(3) $$0 < P(X) < 1 .$$

It guarantees individuals with identical observable characteristics a positive probability to belong both to the participation group and controls (ROSENBAUM & RUBIN, 1983; HECKMAN, LALONDE & SMITH, 1999). Both assumptions together ensure that participation is strongly ignorable and imply that:

(4) $$Y_0, Y_1 \perp Z \mid P(X).$$

As long as outcomes are independent of participation given *X*, then they also do not depend on participation given *P(X)*. Therefore, the multidimensional matching problem is left to a one-dimensional problem. The distribution of potential outcomes will be balanced among participants and counterfactuals (ROSENBAUM & RUBIN, 1983; HECKMAN, ICHIMURA & TODD, 1997, 1998).

A logit model to estimate the propensity score will be applied here, i.e. the probability of participation, given vector *X* containing all observable characteristics, can be defined as:

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

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

After the set up of the underlying assumptions and the prediction of the probability of participation, one parameter that measures the differences in outcome between participants and non-participants is introduced in the next step. In general, the difference in potential outcomes can be captured in the treatment effect for an individual *i*, expressed as follows:

(6) $$TE_i = Y_{i1} - Y_{i0},$$

where $i = 1,\ldots, N$ and *N* represents the total population. Obviously, the individual treatment effect cannot be calculated, because it may not be possible to observe both outcomes for the same agent at the same time. Hence, treatment effects over the average population with counterfactuals for unobserved outcomes need to be derived (HECKMAN et al., 1999). One parameter of interest here is the average treatment effect on the treated (ATT). Applying the composite assumption of "strongly ignorable treatment assignment" (ROSENBAUM & RUBIN, 1983: 43), as expressed in equation (4), the true ATT based on PSM can be written as:

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

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 maximum knowledge score achieved between participants and non-participants that are identical in observable characteristics and adequately weighted by a balanced probability of participation.

In order to obtain matched pairs, CALIENDO AND KOPEINIG (2005) report that there are different matching methods implicating trade-offs in terms of bias and efficiency. Therefore, three different matching estimators are described in the following in order to associate the outcome of participating units to the outcome of their controls. To begin with the most straightforward method, nearest neighbour matching (NNM) implicates to select the non-participant with the smallest distance in propensity score to the participant's propensity score. The nearest neighbour will be matched only once without replacement. This one-to-one matching will cause no concern, as long as the distribution of propensity score of both groups is similar. However, provided that the nearest neighbour is far away, poor matches will be obtained. The average outcome of the matched control will be equally weighted. Hence, the impact estimator is the average difference in knowledge score between participants and controls (SMITH & TODD, 2005). Secondly, radius matching (RM) involves all neighbours within a maximum propensity score distance (caliper), a priori defined, and thus corresponds to the common support assumption. Additionally, poor matches through too distant neighbours are avoided (DEHEJIA & WAHBA, 2002; SMITH & TODD, 2005). Thirdly, Heckman et al. (1997, 1998) recommend kernel-based matching (KM), a non-parametric matching estimator that includes all individuals of the underlying sample of non-participants and weights more distant observed characteristics among both groups down. Hence, kernel-based matching on all control units indicate a lower variance, nevertheless poorer matches may be obtained (CALIENDO & KOPEINIG, 2005). The kernel-based estimator of the ATT describes the mean difference in outcomes while the matched outcome is given by a kernel-weighted average of outcomes of all non-participating units.

Finally, in consideration of the non-randomised selection of farmers in the trypanocide resistance study, it might be possible that unobservable factors like farmers' intrinsic motivation, specific abilities as well as preferences had affected the participation decision. ROSENBAUM (2002) suggests solving this problem of hidden bias by the following bounding approach. Therefore, within the logit model to estimate propensity score (equation 5) the probability of participation *F(·)* needs to be complemented by a vector *U* containing all unobservable variables and their effects on the probability of participation captured by γ:

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

Rearranging the odds ratio of two individuals (*m* and *n*) who are identical in observable characteristics *X*, the resulting relative odds of participation is given by:

(9) $$\frac{\dfrac{P(X_m)}{1 - P(X_m)}}{\dfrac{P(X_n)}{1 - P(X_n)}} = \frac{e^{X_m\beta + U_m\gamma}}{e^{X_n\beta + U_n\gamma}} = e^{[\gamma(U_m - U_n)]}.$$

As long as there is no difference in *U* between the two individuals or if the unobserved variables got no influence on the probability of participation, the relative odds ratio becomes one and the selection process is random. Sensitivity analysis examines now 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). ROSENBAUM (2002) suggests implying the following bounds on the odds ratio of the propensity scores of both individuals:

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

Both individuals have the same probability of participation, provided that they are identical in *X*, only if $e^{\gamma} = 1$. Consequently there will be no selection bias on unobservable covariates. If $e^{\gamma} = 2$, one of the matched individuals may be twice as likely to participate as the other agent (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).

Hence, unbiased impact estimates of a quasi-experimental study design can be obtained in three steps: (i) chose a binary response model with appropriate observable characteristics to predict the probability of participation; (ii) estimate the performance difference between treatment and control group according to selected matching methods that minimise the difference in observables of both groups; and (iii) analyse the effect of unobservable influences on the inference about impact estimates. Based on the implementation of these steps, the following results can be obtained.

4 Results

In order to measure the impact of livestock research on farmers' knowledge the outcome variable is grouped into four different categories:

1) Knowledge about trypanosomosis itself comprising signs, causes, animal re-infection and animals' susceptibility to the disease (maximum score: 26).

2) Curative knowledge and actual control actions in case of trypanosomosis' occurrence including the quality and quantity of trypanocides' use (maximum score: 24).
3) Preventive knowledge and actual preventive strategies applied involving also cattle husbandry and medical management comprising expiry, storage and source of medicines (maximum score: 38).
4) Finally the total knowledge score sums all points from the three categories above (maximum score: 88).

All four knowledge categories are calculated in percentage of maximum score.

In accordance to chosen characteristics that capture all observable relevant differences between participants and non-participants, Table 1 reports the results from the logit model, while the estimated coefficients are expressed in terms of odds of *Z=1*. The summary statistics in Table 1 show that the model is statistically significant. The goodness of fit test achieves a Pearson Chi-square with a high probability value. Hence, the chosen observable characteristics adequately explain the probability of participation. Additionally, the proportion of the total number of predictions that were correctly estimated is about 68.50%.

Table 1: Logit model to predict probability of participation

Dependent variable: Participation **Covariates**			**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
Herd size			1.012**	0.004**
Farming experience of household head			1.843***	0.147***
Number of means of transport			1.043	0.01
Perception of resistance dummy (1 = Resistance)			2.264***	0.182***
Perception of disease dummy (1 = AAT)			1.256	0.054
Country dummy (1 = Burkina Faso)			0.208***	-0.371***
Summary statistics				
Observations	508			
Log-Likelihood	-295.82584			
Pearson Chi2(495)	520.28	Prob>Chi2	0.2086	
Pseudo R-squared	0.142			
Accuracy	68.5%			

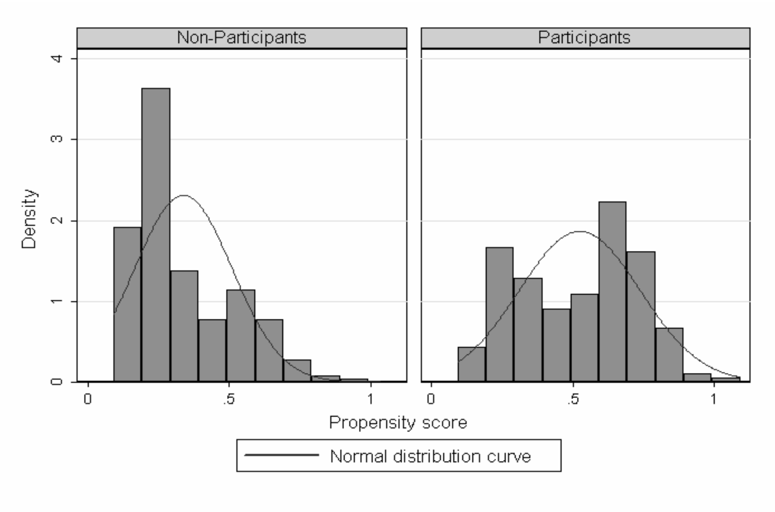
Note: *p<0.1, **p<0.05, ***p<0.01.

Source: own survey

Examining single observables, it is shown that especially farm and village characteristics are significant in the participation model. Each increase in herd size by one cattle is associated with a 1.2% increase in odds of participation. Considering a marginal change in number of cattle, the probability of participation would increase by 0.4%. Farming experience yields even a higher impact as long as more experience both in crop and livestock production enhances the probability of participation about 14.7%. Likewise, the propensity score is increasing by 5.4% when farmers observe their cattle falling sick with AAT. In case farmers perceive ineffectiveness of treatment, which is meant to indicate resistance, the probability of participation is affected even stronger. Moreover, the propensity score is about 37% higher for individuals living in Mali than for Burkinabes.

The resulting predicted probability of participation is plotted in Figure 1. While the propensity score is more equally distributed for participants, the probability distribution of non-participants is skewed to the right. In other words, there are more non-participants than participants with a probability of participation smaller than 50%. Therefore, the application of the common support condition (assumption 3) will be essential for impact estimations.

Figure 1: Histogram of propensity score for non-participants and participants



Source: own survey

Based on the predicted propensity score an appropriate counterfactual group that is as similar as possible to the participating group is matched now. Table 2 shows the impact estimators obtained from the three different matching algorithms. Ensuring that observations are ordered randomly and that there are no large disparities in the distribution of propensity score (Figure 1), one-to-one matching yields the highest and most significant average treatment effects on the treated in all four outcome categories. The nearest neighbour estimate of the average total knowledge gain due to participation is about 3.16%. Since this method produces relative poor matches due to the limitation of information, the attention should be focused on the other two matching algorithms. Here, the estimated impacts of participation in research activities on knowledge score are lower regarding the respective categories.

Following the radius matching algorithm, considering only all neighbours within a caliper of 0.01, the difference in total knowledge scores in percentage of maximum score achieved, is about 2.73%. Moreover, the estimated treatment effect in the category of curative control knowledge and action even accounts for 3.9% at a significance level of 1%. Weighting the average outcome of the matched control with a biweight kernel function and a smoothing parameter of 0.06, like recommended by Silverman (1986), produces also the highest impact estimate due to participation in the category of curative know-how and actual executed control strategies. Similarly to the radius matching estimator in the total score category the kernel-based matching algorithm produces a significant average treatment effect on the treated of 2.78% at the 1% level.

Table 2: Estimated impact of livestock research activities on farmers' knowledge using different matching algorithms

	Knowledge score in % of high scores of		Average treatment effect on the treated
	Participants	**Non-participants**	
Nearest neighbour matching	Using the single closest neighbour		
Knowledge score on disease	25.3	22.93	2.37***
			(4.4)
Knowledge score on control	23.54	19.29	4.25***
			(4.93)
Knowledge score on prevention	16.01	13.0	3.01***
			(5.56)
Total knowledge score	20.81	17.65	3.16***
			(6.59)
Observations	211	211	
Radius matching	Using all neighbours within a caliper of 0.01		
Knowledge score on disease	25.04	23.22	1.82**
			(2.22)
Knowledge score on control	23.17	19.27	3.9***
			(3.17)
Knowledge score on prevention	15.79	13.18	2.6***
			(3.44)
Total knowledge score	20.54	17.81	2.73***
			(4.03)
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**
			(2.36)
Knowledge score on control	23.55	19.91	3.64***
			(3.02)
Knowledge score on prevention	16.03	13.18	2.85***
			(3.77)
Total knowledge score	20.81	18.03	2.78***
			(4.15)
Observations	210	293	

Note: T-statistics in parentheses and *p<0.1, **p<0.05 and ***p<0.01.

Source: own survey

Consequently, it can be proven that livestock research activities generate in fact a significant gain in farmers' knowledge on trypanosomosis and improve both curative and preventive strategies.

Following ROSENBAUM (2002), an appropriate control strategy of hidden bias is to examine the sensitivity of significance levels. Here, for several values of e^{γ} bounds are calculated on the significance level, and hence, the null hypothesis of no effect of participation on potential outcomes, respectively on knowledge score is then tested. Therefore, the question arises at which critical impact level of the unobservable the inference about the treatment effect on knowledge will be undermined, as indicated by the loss of significance (DIPRETE & GANGL, 2004). Table 3 compares the sensitivity of treatment effects on different knowledge scores using the three introduced matching algorithms.

Table 3: Sensitivity analysis with Rosenbaum's bounds on probability values

	Upper bounds on the significance level for different values of e^{γ}				
	$e^{\gamma}=1$	$e^{\gamma}=1.25$	$e^{\gamma}=1.5$	$e^{\gamma}=1.75$	$e^{\gamma}=2$
Nearest neighbour matching	Using the single closest neighbour				
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 neighbours 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

Overall, robustness results produced by Rosenbaum's bounds are quite similar. However, kernel-based matching produces the most robust treatment effect estimates with respect to hidden bias especially in the category of preventive knowledge and action as well as in the fourth class were all points are summarised. Matched pairs might differ up to 100% ($e^{\gamma}=2$) in unobservable characteristics, while the impact of participation on preventive knowledge as well as on total knowledge would be still significant at a level of 5% (*p*-value = 0.023 and *p*-value = 0.0144, respectively). The same knowledge categories are robust to hidden bias up to an influence of $e^{\gamma}=2$ at a significance level of 10% following the radius matching approach. Also the less qualified matching algorithm of nearest neighbour matching is robust to selection bias on unobservable characteristics up to an impact level of $e^{\gamma}=2$ in the fourth category. The estimated treatment effects on knowledge about trypanosomosis itself as well as on the category of curative knowledge and action are sensitive to hidden bias, at a smaller unobservable impact level of $e^{\gamma}=1.5$.

Nevertheless, it has to be considered that these sensitivity results are worst-case scenarios, although they indicate information about uncertainty within the matching estimators of treatment effects (ROSENBAUM, 2002).

5 Conclusions

Propensity score matching (PSM) allows measuring the short-term impact of a natural resource management project on farmers' knowledge and practice of trypanosomosis control. Due to the quasi-experimental design of the intervention, with non-randomised selection of villages and farmers, PSM is effective to overcome the selection bias on observable characteristics of project participants and non-participants. PSM creates then reliable impact estimates, respectively treatment effects, when the predicted probability of participation given observable treatment-independent covariates is balanced among those who are identical in these observables. Hence, matched participants and non-participants can only be distinguished by their treatment attribute and unbiased performance differences can be obtained.
Using three different matching algorithms significant and robust differences between matched participants and non-participants regarding cattle farmers' knowledge were identified. Hence, it can be concluded that the gain in farmers' knowledge is attributable directly to participation in the research intervention. The strongest effect of the research intervention is on the curative knowledge of trypanosomosis and subsequent adequate control decisions. Moreover, significant advancements in preventive strategies are also observable. Overall, the research project has been effective to increase farmers' knowledge and to improve their practices. However further research is needed to establish the efficiency of the research investment.

References

AAKVIK, A. (2001): Bounding a Matching Estimator: The Case of a Norwegian Training Program. In: Oxford Bulletin of Economics and Statistics, 63: 115-143.

AFFOGNON, H. (2007): Economic Analysis of Trypanocide Use in Villages under Risk of Drug Resistance in West Africa. Doctoral Dissertation, Leibniz Universität Hannover, Germany.

BAKER, J.K. (2000): Evaluating the Impact of Development Projects on Poverty: a Handbook for Practioners. World Bank, Washington DC.

BUDD, L.T. (1999): DFID-Funded Tsetse and Trypanosomiasis Research and Development since 1980. Volume 2. Department for International Development (DFID), London, UK.

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 Labour (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: 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: 271-310.

HECKMAN, J.J., ICHIMURA, H. 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: 605-654.

HECKMAN, J.J., ICHIMURA, H. and P.E. TODD (1998): Matching as an Econometric Evaluation Estimator. In: The Review of Economic Studies, 65, pp. 261-294.

HECKMAN, J.J., LALONDE, R.J. and J.A. SMITH (1999): The Economics and Econometrics of Active Labour Programs. In: O. Ashenfelter & D. Card (Eds.), Handbook of Labor Economics (pp. 1865-2097). The Netherlands: Elsevier.

RAITZER, D.A. and T.G. KELLY (2008): Assessing the Contribution of Impact Assessment to Donor Decisions for International Agricultural Research. In: Research Evaluation, 17: 187-199.

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

ROSENBAUM, P.R. and D.B. RUBIN (1983): The Central Role of Propensity Score in Observational Studies for Causal Effects. In: Biometrika, 70: 41-55.

SILVERMAN, B.W. (1986): Density Estimation for Statistics and Data Analysis. London: Chapman & Hall.

SMITH, J. and P. TODD (2005): Does Matching Overcome LaLonde's Critique of Non-Experimental Estimators? In: Journal of Econometrics, 125: 305-353.

ZILBERMAN D. and H. WAIBEL (2007): International Research on Natural Resource Management: Advances in Impact Assessment. Oxfordshire: CABI.