Adoption of agroforestry-based biofuel systems in South India

F. Dalemans; B. Muys; M. Maertens

KU Leuven, Earth and Environmental Sciences, Belgium

Corresponding author email: floris.dalemans@kuleuven.be

Abstract:

Agroforestry-based biofuel production has recently been proposed as a rural development strategy in the South. However, there exists a complete lack of empirical evidence on farmer adoption rates and determinants for these novel systems. This study describes adoption rates of oilseed tree mixtures on smallholder farms in Hassan district, South India, and quantifies how these rates are determined by a biofuel extension program (BP) and farm(er) characteristics. This is done through a set of regression-based analyses, addressing various forms of selection bias. The findings reveal that although 60% of the farmers cultivate oilseed trees, oilseed collection rates are generally low (13%), and the adoption of both practices is driven by different determinants. More specifically, BP activities are found to stimulate tree cultivation and therefore agroforestry establishment, but not seed collection and biofuel production. This calls for a better understanding of adoption profitability in function of the opportunity costs of land, labour and capital involved, and conditional on these results for intensifying BP activities and value chain development.

Acknowledegment: Thanks go to the UAS Bangalore Department of Agricultural Extension and all individual enumerators for their assistance in data collection and data entry. We would also like to thank the Biofuel Park Hassan staff for their scientific support and for their assistance in data collection. This work was supported by the Research Foundation Flanders (FWO) [Aspirant PhD grant].

JEL Codes: Q23, Q42

#495

Adoption of agroforestry-based biofuel systems in South India

Abstract

Agroforestry-based biofuel production has recently been proposed as a rural development strategy in the South. However, there exists a complete lack of empirical evidence on farmer adoption rates and determinants for these novel systems. This study describes adoption rates of oilseed tree mixtures on smallholder farms in Hassan district, South India, and quantifies how these rates are determined by a biofuel extension program (BP) and farm(er) characteristics. This is done through a set of regression-based analyses, addressing various forms of selection bias. The findings reveal that although 60% of the farmers cultivate oilseed trees, oilseed collection rates are generally low (13%), and the adoption of both practices is driven by different determinants. More specifically, BP activities are found to stimulate tree cultivation and therefore agroforestry establishment, but not seed collection and biofuel production. This calls for a better understanding of adoption profitability in function of the opportunity costs of land, labour and capital involved, and conditional on these results for intensifying BP activities and value chain development.

Keywords: biofuel, agroforestry, farmer adoption, extension, econometric analysis

1. Introduction

Liquid biofuel production has recently been proposed as a strategy for rural development in developing countries (Demirbas and Demirbas, 2007). Smallholder farming communities could act as feedstock producers (and processors) in biofuel value chains, thereby generating income, employment, trade and energy supply at a local level (Ewing and Msangi, 2009). In particular, much hope is set on ‘alternative’ biofuel systems, which aim at minimizing the interference of biofuels with food production and markets (Ewing and Msangi, 2009). This is a prerequisite for any biofuel program in India (Biswas and Pohit, 2013), which has the highest absolute rate of undernourishment in the world (FAO et al., 2017). At the same time, India is depending for 70% of its oil demand on imports; a share which is projected to surpass 90% by 2040 (IEA, 2015). This has resulted in various government policies promoting widespread cultivation of non-edible oilseed trees for biodiesel production on marginal lands (Biswas and Pohit, 2013). Initial programs were mainly based on large-scale plantations of jatropha (*Jatropha curcas*) on wastelands, but seed yields proved to be limited and highly variable under low input regimes, resulting in economic unviability and limited pro-poor potential (Achten et al., 2014; van Eijck et al., 2014). Therefore, alternative biofuel approaches are being explored.

Small-scale agroforestry-based biofuel production for local use is being suggested as a possible solution (Achten et al., 2014; Sharma et al., 2016). This approach carefully integrates multipurpose oilseed trees within the existing farming system. While oilseed yields might be also limited in such low input – high diversity – high resilience systems (Tilman et al., 2006), the trees bring along multiple other products, uses and co-benefits, which add to the viability of the approach. Agroforestry set-ups are known to introduce various co-benefits, including soil conservation and fertility increase, reduced disease pressure, carbon sequestration, shading and hedging, labour and income diversification, risk mitigation, and increased farm resilience (Sileshi et al., 2007).

The rationale of agroforestry-based biofuel systems is thus based on principles widely supported by academic literature. However, Pattanayak et al. (2003) and Mercer (2004) note that no matter how efficient, productive and/or innovative agroforestry projects may be, low adoption rates and rapid disadoption widely occur, while substantial impacts may only materialize if these systems are adopted and maintained over a long time. This calls for a thorough understanding of how farmers perceive and adopt these novel systems, as well as for identifying and quantifying (dis)adoption determinants (Mercer, 2004). On the one hand, program design (e.g., input support, production system, value chain organisation), extension (e.g., mode, duration, intensity) and diffusion will have a crucial role to play in stimulating adoption (Dalemans et al., in press; Glendinning et al., 2001; Pattanayak et al., 2003). On the other hand, a wide range of farm and farmer characteristics may influence adoption decisions, see for instance the seminal work of Feder et al. (1985) and the meta-analysis by Pattanayak et al. (2003). Although these articles survey the large existing empirical adoption literature, only few studies exist on biofuel adoption, e.g., Axelsson et al. (2012), Basinger et al. (2012) and Soto et al. (2015). In addition, existing studies usually focus on jatropha and monoculture set-ups (plantations). Therefore, there exists a complete lack of empirical evidence on farmer adoption of agroforestry-based biofuel systems in developing countries.

An example of an agroforestry-based biofuel approach is being implemented since 2007 in Hassan district, Karnataka state, India, by a university – government partnership[1]. This program (hereafter referred to as Biofuel Program or BP) aims to establish oilseed tree mixtures in smallholder farms on borders, in homegardens, and on fallow land, and on communal lands. It mainly focuses on native species, which are adapted and non-invasive to local ecosystems, while local communities are familiar with their various uses. The principal species are pongamia (*Millettia pinnata*), neem (*Azadirachta indica*), mahua (*Madhuca indica*), simarouba (*Simarouba glauca*) and jatropha (*Jatropha curcas*). For pongamia and neem in particular there exist many traditional uses of wood, leaves, fruits, seed cake[2] and seed oil. While seed oil can be blended in diesel or processed to biodiesel, it is traditionally known as a lamp fuel and for its pesticidal, medicinal and industrial applications. Corresponding value chains exist, and oilseed collection on private and communal land is traditionally known as a marginal activity (Altenburg et al., 2009). The Biofuel Program primarily focuses on extension through conducting awareness campaigns and trainings on oilseed tree cultivation and biofuels. Furthermore, it distributes high-yielding tree seedlings among farmers, and establishes biodiesel cooperatives within villages to streamline activities. The Biofuel Program also hands out small-scale oil-expelling equipment, and provides various kinds of market support. So far, the program has only been evaluated in terms of provided extension and support activities. There is very limited insight on actual adoption rates, and how they are impacted by the Biofuel Program.

The aim of this paper is to investigate farmer perception and adoption of oilseed trees in Hassan district, and identify and quantify adoption determinants, through econometric analyses of cross-sectional survey data collected in 2015. Thereby it contributes to the empirical evidence on farmer adoption of biofuel systems, in particular for agroforestry-based approaches using novel oilseed species. The analysis will focus on the role of extension, program implementation and farm(er) characteristics as drivers and barriers to adoption. With regard to the effect of program design on adoption, the reader is further referred to Dalemans et al. (in press).

[1] The Karnataka State Bio Energy Development Board and the University of Agricultural Sciences Bangalore
[2] Seed cake is a by-product of seed oil extraction, and has pesticidal, fertilizer and fodder applications.

2. Methodology and Data

2.1. Research area

Hassan district has a population of about 1.8 million inhabitants, 79% of whom live in rural areas. It is constituted by about 2600 villages, which are clustered into 38 administrative units termed hoblis (DCO, 2014). Its geophysical diversity is prevalent from the three agro-ecological zones (dry, transition, hill) characterized by a distinct rainfall gradient (Figure 1). Correspondingly, a wide variety of crops are cultivated, including various plantation crops such as coconut in the dry and transition zone, and coffee, pepper and cardamom in the hill zone. Most farmers are smallholders: 65% of the landholdings are smaller than 1 hectare, 89% smaller than 2 hectares (DAC&FW, 2017).

2.2.Biofuel Program (BP) extension and implementation

The Biofuel Program is providing extension and support on oilseed trees in Hassan district since 2007. This comprises a range of activities. First, the BP conducts biofuel awareness campaigns and trainings. These information and demonstration sessions focus on the cultivation and uses of various oilseed tree species, and in particular on the biofuel value chain. They generally take place on-site in local villages, and by 2015 at least one training had been conducted in half of the villages within Hassan district. Furthermore, 290 sessions had been organized in the central extension centre for visiting groups. Second, the BP organizes planting programs throughout the district. By 2015, 1.55 million seedlings of various oilseed tree species had been distributed for planting on borders, in homegardens, on fallow land and on communal lands. Third, the BP has supported 470 villages by 2015 in establishing biodiesel cooperatives. These cooperatives should take the lead in coordinating biofuel activities in the community (seed collection, storage, marketing, processing and use), and act as a meeting point for local farmers and the extension agency. Fourth, the BP aims to reinforce the forward linkages of the biofuel value chain in two ways. On the one hand, they provide marketing support through minimum oilseed support prices, on-farm pick-up, local processing of oilseed to biofuels, and redistribution of the end products to local consumers. On the other hand, they have distributed a small-scale oil expeller in 19 communities which expressed profound interest in the biofuel value chain, to enable on-site extraction and use of seed oil.

2.3. Data collection

To investigate farmer perception and adoption of oilseed trees in Hassan district, cross-sectional household survey data were collected in Hassan district in the period August – September 2015. A three-stage stratified random sampling was used. In the first stage, 1, 2 and 3 specific hoblis were selected in the hill, dry and transition zone, respectively[3] (Table 1). Purposeful selection was done to maximize the variation in BP extension and implementation within hoblis. Correspondingly, in the second stage the villages in each hobli were separated in four BP extension and implementation strata[4]:

(1) villages where no BP activities have taken place ("control - C");
(2) villages where awareness campaigns and trainings have taken place, but no cooperative has been formed and no oil expeller has been distributed ("training - T")
(3) villages where awareness campaigns and trainings have taken place, a cooperative has been formed, but no oil expeller has been distributed ("cooperative" or "association" - A")
(4) villages where awareness campaigns and trainings have taken place, a cooperative has been formed, and an oil expeller has been distributed ("expeller - E")

Villages were randomly drawn within these strata resulting in 6 sampled villages for each hobli (Table 1; Figure 1). Finally, in the third stage 11 farm-households were drawn in each of the 36 villages through systematic sampling, which resulted in a final sample of 396 households. For the analysis in this paper, 6 households are dropped due to incomplete and/or erroneous data, which reduces the final sample to 390 households.

Household survey data were collected using a quantitative structured questionnaire. The questionnaire included an extensive module to assess farmers' knowledge, perception and adoption of five principal oilseed trees (pongamia, neem, mahua, simarouba and jatropha), and their participation in the biofuel value chain. In-depth data on farmers' involvement in various BP activities were collected. This was supplemented with data on a wide range of farm(er)

---

[3] The number of selected hoblis for each zone reflects its share in the district population.

[4] Planting programs have been implemented in all A- and E-villages, some of the T-villages, and none of the C-villages.

characteristics through detailed modules on household demographics, land and non-land assets, farm production and marketing, employment and income, and social network.

Table 1. Characterization of the sampled hoblis (first stage) and distribution of the sampled villages (second stage). Adapted from (Dalemans et al., in press).

Agro-ecological zone	Hobli	Annual rainfall (mm)	Amount of sampled villages by stratum				Total per hobli
			Control (C)	Training (T)	Cooperative (A)	Expeller (E)	
Dry	Nuggehalli	613	3	1	1	1	**6**
	Arsikere	709	0[a]	2	1	3	**6**
Transition	Halekote	737	1	1	1	3	**6**
	Kattaya	891	2	1	1	2	**6**
	Palya	1258	2	2	1	1	**6**
Hill	Hetthur	2477	2	2	2	0[a]	**6**
Total per village type			**10**	**9**	**7**	**10**	

[a] A zero value indicates that there are no villages within this stratum for the corresponding hobli.

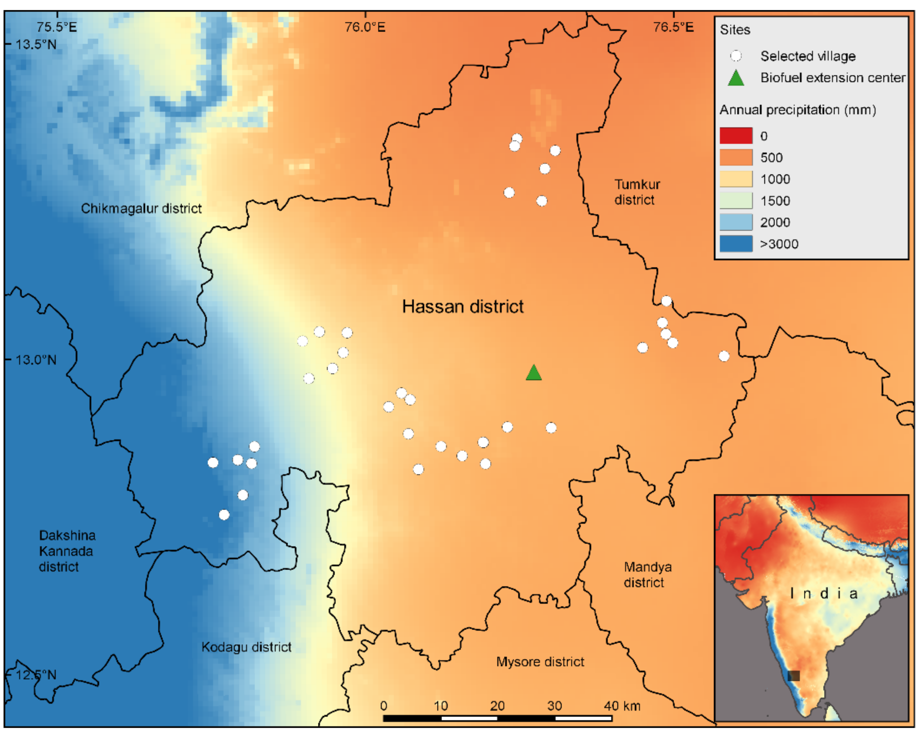

Figure 1. Map of Hassan district, Karnataka state, India. The map locates the biofuel extension centre, sampled villages and rainfall gradient. Reproduced from Dalemans et al. (in press).

2.4. Econometric framework

To analyze the effect of BP activities and farm(er) characteristics on adoption using cross-sectional data, a variety of regression models are estimated. Before specifying these, further clarification is required on the conceptualization of adoption, and on the endogeneity of the BP treatment variables.

2.4.1. The complex nature of adoption: sample selection

In most empirical work adoption is conceived as the take-up of a specific technology or practice, and correspondingly considered as a single binary choice. In this case study however the adoption concept should be considered more ambiguously.

First, an explicit distinction should be made between the adoption of two components: tree cultivation (defined as the presence of at least one oilseed tree species (pongamia, neem, mahua, simarouba or jatropha) on the household's landholding) and seed collection (defined as seed collection by the household from at least one of the aforementioned species in the past 12 months). This is on the one hand because it takes several years for these species to start yielding[5], implying that some farmers might be cultivating trees, yet they are not able to collect seeds from these. On the other hand these species have multiple uses and co-benefits, implying that some farmers might be cultivating the trees but not for oilseed collection. To put it another way, these arguments imply that the possibility for oilseed collection is conditional upon having mature trees, but not a necessary consequence of it. More technically, adoption can be considered a two-stage sequential process, in which the seed collection adoption choice (stage two, y_s) is unobserved for farmers who do not have mature trees (stage one, y_m). In terms of a generalized linear model:

$$\text{Selection equation:} \quad y_m^* = g(X_m'\beta_m + T_m'\delta_m + \varepsilon_m) \quad (1)$$

$$\text{Structural equation:} \quad y_s^* = g(X_s'\beta_s + T_s'\delta_s + \varepsilon_s) \quad (2)$$

[5] The length of this maturation period varies among species: on average 3 years for jatropha, 4 years for simarouba, 5 years for pongamia and neem, 9 years for mahua.

Where y_m^* and y_s^* are latent adoption probabilities, X_m and X_s are control variable sets, T_m and T_s are Biofuel Program treatment variable sets, β_m, β_s, δ_m, δ_s are coefficient sets, ε_m and ε_s are error terms, g( ) is a link function, and with observation rule:

$$\text{if } y_m^* > 0: \quad y_m = 1 \text{ and } \begin{cases} \text{if } y_s^* > 0: & y_s = 1 \\ \text{if } y_s^* \leq 0: & y_s = 0 \end{cases}$$

$$\text{if } y_m^* \leq 0: \quad y_m = 0 \text{ and } \quad y_s \text{ is not observed}$$

This interpretation raises the question whether the subsample of farmers having mature trees ($y_m = 1$) is random. One could reasonably argument that farmers who are willing to collect oilseeds put on average more effort into cultivating trees to this end. This would lead to a sample selection bias when estimating the structural equation separately on the subsample. The Heckman correction (Heckman, 1976) can be used to account for this sample selection bias. This usually requires at least one selection variable which determines tree maturity in equation 1, but not seed collection conditional on having mature trees in equation 2. The following variables are proposed because they relate to the natural occurrence of these species in the area: a respondent-reported variable indicating whether wild oilseed trees occur naturally on their land ("Natural tree occurrence"), and the amount of parcels in the farmer's landholding ("Amount of parcels")[6].

Second, most oilseed tree species are native and wild trees are prevalent in the district. This might enable farmers to collect seeds from communal lands and forests[7]. Furthermore, it could well be that some farmers are allowed to collect seeds from other farmers' private land. These arguments question the proposition above that seed collection is strictly conditional upon having mature trees on one's own land. They call for estimating the structural equation (= seed collection adoption) on the entire sample, using tree maturity (y_m) as a regressor rather than as a selection indicator. Sample selection bias is then no issue anymore, which raises the question whether the selection equation (= having mature trees) is still relevant to estimate on its own. From a long-term policy viewpoint it would be more useful to assess the determinants of tree cultivation adoption (y_t) in general,

[6] One would assume that both natural tree occurrence and amount of parcels – when controlling for total landholding size – increase the probability of having trees – and by extension mature trees – on the farm, but that for the subsample of farmers having mature trees, it does not determine the probability of seed collection.
[7] In particular because the BP planting programs are also targeting communal lands.

whether these trees are already mature or not. Thus, in terms of a generalized linear model, the following equations are estimated independently and both on the full sample:

$$\text{Tree cultivation equation:} \quad y_t^* = g(X_t'\beta_t + T_t'\delta_t + \varepsilon_t) \qquad (3)$$

$$\text{Seed collection equation:} \quad y_s^* = g(X_s'\beta_s + T_s'\delta_s + \varepsilon_s) \qquad (4)$$

Where y_t^* and y_s^* are latent adoption probabilities, X_t and X_s are control variable sets, T_t and T_s are Biofuel Program treatment variable sets, β_t, β_s, δ_t, δ_s are coefficient sets, ε_t and ε_s are error terms, g( ) is a link function.

2.4.2. The BP treatment effect: self-selection and simultaneity

This study aims to assess the effect of three BP activities on tree cultivation and seed collection adoption: awareness campaigns and trainings, biodiesel cooperative formation and expeller distribution[8].

The first is considered as a binary variable indicating whether at least one household member has participated in at least one session on-site and/or in the central extension centre ("BP training"). However, participation in these sessions is voluntary, and one could reasonably argument that farmers who are willing to adopt self-select into the treatment. This would lead to a positive self-selection bias. To address this bias in cross-sectional models, instrument-based methods can be used. The following instruments are proposed as they are measurements of social capital and networks, and access to information: a binary variable indicating whether at least one household member is a member of a village organization / group ("Group membership"), and a principal component quantifying household exposure to radio, tv, internet and newspapers ("Media exposure") (larger values indicate higher exposure). These variables are expected to positively influence the probability of program participation (= relevant), but not to have a direct effect on the probability of tree cultivation and seed collection conditional on the set of control variables (= exogenous).

[8] The other two BP activities, i.e., planting programs and oilseed marketing support, are not considered here. This is because farmer participation in a planting program is a result, rather than a predictor, of willingness to adopt tree cultivation. Receiving oilseed marketing support is similarly a result, rather than a predictor, of oilseed collection adoption.

The latter two BP activities are considered as binary variables indicating respectively whether a biodiesel cooperative has been formed in the village ("BP association"), and whether an expeller has been distributed to the village ("BP expeller"). These two indicators are particularly likely to be endogenous, because these BP interventions are implemented in a later stage only, and reserved for villages which haven proven (extensive) interest in tree cultivation and oilseed collection, in particular for "BP expeller". Causality might therefore run in two directions: presence of a cooperative and/or expeller might not only influence adoption, but is also a result of it. This would result into a positive simultaneity bias. Since the dependent variable (adoption) is the exact criterion for assigning the treatment, it is extremely hard to address this bias in a cross-sectional analysis. Therefore, two strategies are considered. First, village fixed effects ('cluster fixed effects' (Imbens and Wooldridge, 2007)) can be added to the regression model, since this removes variables which have no within-village variation, like the cooperative and expeller indicators, from the model. While this removes the bias associated with these indicators, it obviously also implies that their effect is not modelled anymore. Second, a model can be run without bias correction, for which the estimated coefficients can be assessed in combination with the hypothesized direction of bias. However, obtaining the direction of bias is in general complicated in models with many correlated regressors (Wooldridge, 2010).

2.4.3. Regression models

As adoption is a binary dependent variable, regression approaches are mostly based on probit models. However, the issues raised in sections 2.4.1. and 2.4.2. pose difficulties to this approach. First, correcting for self-selection bias is difficult in nonlinear models, especially if the endogenous variable is also discrete (Imbens and Wooldridge, 2007; Wooldridge, 2014). Bivariate probit models can be applied in such cases (Greene, 2012). Second, while sample selection bias can be addressed with a sample selection probit model (Van de Ven and Van Praag, 1981), correcting for both sample selection and self-selection bias at the same time is difficult when the outcome variable is binary (Wooldridge, 2010). Last, cluster fixed effects should not be applied in a probit model due to the incidental parameters problem (Greene, 2012; Wooldridge, 2010). In addition, if there is no variation in the dependent variable within a cluster, these observations will be dropped in the estimation when using cluster fixed effects.

Linear probability models on the other hand can address these issues more flexibly. Therefore, a stepwise approach is used. The first step only focuses on the role of sample selection bias, and estimates equations 1 and 2 using the sample selection probit model (Heckman, 1976; Van de Ven and Van Praag, 1981)[9]. Village fixed effects are not added. In the second step, equations 3 and 4 are estimated with a linear probability model to investigate the effect of BP training[10]. To account for self-selection bias, the control function approach (or alternatively, two-stage residual inclusion) to instrumental variable estimation is applied because of two advantages over the two-stage least squares approach (Wooldridge, 2015). First, it allows for a straightforward robust test of BP training exogeneity. Second, it flexibly allows to interact control variables with the endogenous variable. Village fixed effects are added to remove simultaneity bias due to BP village-level activities. In a final step, equations 3 and 4 are estimated with hobli[11] instead of village fixed effects, to assess the effects of BP association and BP expeller, acknowledging the simultaneity bias.

In all models a set of control variables (X_m, X_s, X_t) is included as well, to identify and quantify farm(er) characteristics acting as drivers and barriers to adoption. Human capital is quantified by indicators for the gender, age and educational level[12] of the household head, and by the amount of labourers and dependents in the household. Physical capital is measured by the total landholding size. Wealth is measured by a principal component based index of household assets and living standards (higher values indicate more assets / higher living standards). Also amount of tropical livestock units is used as a control variable. Tree ownership in 2007 is reconstructed based on recall data.

Finally, the clustered sampling approach should be taken into account when calculating the standard errors. While the use of cluster fixed effects may substantially decrease the possibility of within-cluster error correlation, it does not necessarily remove it entirely (Cameron and Miller, 2015). Therefore, even though heteroscedasticity-robust standard errors will be reported, these might still be biased. Since the amount of clusters is relatively low (6 if clustered on hobli level, 36 if clustered on village level), it is not advisable to use cluster-robust standard errors, which are

[9] For the probit model, the link function in equations 1 and 2 is the normal cumulative distribution function.
[10] For the linear probability model, the link function in equations 3 and 4 is the identity function.
[11] A hobli is an administrative unit in Hassan district, see Section 2.3.
[12] Indicator which equals one if the household head completed at least 8 years of schooling.

only asymptotically valid (Cameron and Miller, 2015). Therefore, wild cluster bootstrapped t-statistics (with 100,000 residual resamples) are used to account for cluster effects (Cameron and Miller, 2015) and reported (= p-btrap) along with the p-values based on heteroscedasticity-robust standard errors (= p).

3. Results and Discussion

3.1. Descriptive statistics: adoption and BP participation

The majority of farmers are cultivating oilseed trees on their land (60%), although for 4 out of 10 farmers these trees are still immature (Table 2). However, seed collection rates are substantially lower: only 13% of all farmers is collecting oilseeds. This figure increases to 31% when only considering farmers having mature trees. On the one hand, this indicates that tree maturity is likely to be a strong determinant for seed collection (in eq. 4), although it is not a strict condition (in eq. 2): 2.4% of the farmers not having mature trees do collect oilseeds. On the other hand, still 2 out of 3 farmers who have mature trees are not collecting the oilseeds. This supports the case that was made in Section 2.4.1 for not considering adoption as a single binary choice, but as two separate stages: adoption of tree cultivation and adoption of seed collection.

Next, when considering these figures across villages stratified by degree of BP implementation, tree cultivation rates – and correspondingly tree maturity rates – are only substantially higher in expeller villages (Table 2). However, this does not apply to the seed collection rates, in particular when controlling for tree maturity. Remarkably, these seed collection rates are substantially lower in cooperative villages. This corresponds to the striking observation that none of the households in these cooperative villages are aware of the cooperative, and consequently also not a member of it (Table 2). Further consideration of BP participation rates reveals that only 18% of the households have attended at least one BP awareness meeting and training (Table 2). This figure is substantially higher in expeller villages, while even few households in the control villages have attended sessions in other villages and/or at the extension centre. In contrast to the cooperative villages, in the expeller villages the cooperatives are more firmly established: 62% of the sampled households know them, 26% are a member. This does not apply to the use of the oil expeller, which is only done by one household in the entire sample. In addition, out of the fifty households collecting oilseeds, only one household sells these to the Biofuel Park, while the others sell them to traders, markets or other persons.

Finally, in table 3 adoption rates are compared between households who have attended at least one BP awareness meeting and training (18%), and those who have not (72%). Tree cultivation rates – and correspondingly tree maturity rates – are significantly higher for the former. This seems to be the main driving factor for the significantly higher unconditional seed collection rates for trained households; when controlling for tree maturity, the difference is still substantial but not significant anymore (p = 0.215).

3.2. Adoption of seed collection, conditional on tree maturity

The sample selection probit model estimates for seed collection probabilities conditional on tree maturity, are presented in Table 4. The null hypothesis that there is no correlation between the error terms ε_m and ε_s , is not rejected (p = 0.140). In other words, the estimation results do not indicate that a sample selection bias is present. Also, note that the proposed selection variables "Natural tree occurrence" and "Amount of parcels" do not prove to be suitable: the former has a strongly significant negative effect in the structural equation, while the latter is not significant in the selection equation. However, other variables could serve as selection variables, including "Amount of livestock units" and "BP training". The latter corresponds to the descriptive results (Table 3).

The sample selection probit model was conceived based on the propositions that seed collection is strictly conditional on tree maturity, and that the seed collection equation independently estimated on the subsample introduces bias. The results of Sections 3.1 and 3.2 support neither of these statements. This implies not only that both equations can be estimated independently, but that the seed collection equation can (or rather should) be estimated on the full sample. This is exploited to the fullest by switching to the linear probability models (see Section 3.3). First, the statistical power in the seed collection equation strongly increases by extending the amount of observations from 141 to 390. Second, it does not require some of the strong assumptions for consistency made by the sample selection probit model, i.e., bivariate normality and homoscedasticity. Third, it allows to account for the endogeneity of the BP program through instrumenting and village fixed effects. Therefore, results of the sample selection probit model will not be further discussed, also as they are largely consistent with the results of the linear probability models (see Section 3.3).

Table 2. Adoption and Biofuel Program participation rates across villages in the sample.

	Total		Control villages		Training villages		Cooperative villages		Expeller villages	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Number of households	390		111		95		76		108	
Tree cultivation, tree maturity, and seed collection rates										
Tree cultivation	0.60	(0.29)	0.57	(0.33)	0.42	(0.22)	0.56	(0.28)	0.81	(0.18)
Tree maturity	0.36	(0.23)	0.34	(0.28)	0.25	(0.17)	0.37	(0.28)	0.48	(0.17)
Seed collection, unconditional on tree maturity	0.13	(0.13)	0.16	(0.16)	0.11	(0.13)	0.04	(0.05)	0.18	(0.13)
Seed collection, conditional on tree maturity	0.31	(0.29)	0.38	(0.37)	0.36	(0.27)	0.07	(0.08)	0.36	(0.27)
Biofuel Program participation rates										
Awareness meetings and trainings	0.18	(0.21)	0.04	(0.06)	0.11	(0.10)	0.16	(0.23)	0.41	(0.20)
Cooperative knowledge[a]	-		-		-		0.00	(-)	0.62	(0.28)
Cooperative membership[a]	-		-		-		0.00	(-)	0.26	(0.19)
Expeller use	-		-		-		-		0.01	(0.03)

[a] Cooperative knowledge refers to how many households know the biodiesel cooperative in the village, cooperative membership to how many households are member of it.

Table 3. Adoption rates according to whether households attended BP trainings or not.

	Total	Trained households	Non-trained households	
Number of households	390	71	319	
Tree cultivation	0.60	0.80	0.55	***
Tree maturity	0.36	0.52	0.33	***
Seed collection, unconditional on tree maturity	0.13	0.21	0.11	**
Seed collection, conditional on tree maturity	0.31	0.41	0.28	

Note: Fisher's two-sided exact test is used to test differences in means for trained and non-trained households. Significant effects are indicated as follows: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 4. Estimates for the sample selection probit model, with tree maturity (y_m) as selection equation, and oilseed collection (y_s) as structural equation.

	Selection equation – Tree maturity			**Structural equation – Seed collection**		
	β	(SE)	p	β	(SE)	p
Gender of HH head (1 = male)	0.015	(0.122)	0.904	-0.041	(0.136)	0.765
Age of HH head (years)	0.007	(0.023)	0.752	0.002	(0.026)	0.940
Education of HH head	-0.144	(0.233)	0.536	-0.227	(0.271)	0.402
Number of HH dependents	0.005	(0.006)	0.421	-0.005	(0.008)	0.474
(Number of HH dependents)^2	0.237	(0.168)	0.158	-0.330	(0.203)	0.104
Number of HH labourers	0.302	(0.565)	0.593	0.231	(1.165)	0.843
(Number of HH labourers)^2	0.126	(0.267)	0.638	-0.727	(1.220)	0.551
Exploited land (10^{-1} acres)	0.119	(0.267)	0.657	-0.397	(0.305)	0.192
(Exploited land)^2 (10^{-1} acres)^2	-0.006	(0.041)	0.881	0.056	(0.051)	0.274
Amount of tropical livestock units	**0.256**	**(0.117)**	**0.029**	0.228	(0.186)	0.220
(Amount of tropical livestock units)^2	**-0.042**	**(0.022)**	**0.055**	-0.064	(0.048)	0.182
Assets and living standards index	0.001	(0.037)	0.979	**-0.093**	**(0.042)**	**0.028**
Natural tree occurence (1 = yes)	**1.038**	**(0.161)**	**0.000**	**-1.074**	**(0.206)**	**0.000**
Amount of parcels	0.053	(0.069)	0.444	-0.010	(0.080)	0.902
BP training (1 = yes)	**0.350**	**(0.180)**	**0.052**	-0.055	(0.202)	0.786
Constant	-2.170	(0.611)	0.000	2.737	(0.716)	0.000
Village fixed effects		(not included)			(not included)	
Amount of observations		390			141	

3.3. Adoption of tree cultivation & seed collection

3.3.1. Tree cultivation

The tree cultivation equation aims primarily to quantify the effect of BP training on adoption. However, a substantial amount of farmers (42.5%) already cultivated trees in 2007, and many of these trees might still be present today. For these households, BP training will not be able to make a difference, except by influencing the decision whether to cut the trees or not. Therefore, tree ownership in 2007 is added as a regressor to capture time-constant unobserved heterogeneity, but also interacted with BP training. The interaction will prevent the described mechanism from "diluting" the real (or potential) effect of BP training. The linear probability model estimates for this specification, using village fixed effects, are presented in Table 5. The base model estimates reveal significant effects of tree ownership in 2007, BP training and their interaction. As hypothesized, for farmers who did not have trees in 2007, BP training significantly increases the probability of having trees today (main effect: 0.23), while for farmers who already had trees in 2007, BP training has no significant effect on that probability (main + interaction effect: -0.09 ($p = 0.144$)).

Also several control variables prove to be drivers of adoption. Household head schooling could lead to better insight into the benefits of agroforestry-based innovations, and higher ability to realize them (Pattanayak et al., 2003). The positive effect of tropical livestock units (although with decreasing rate) could relate to the fodder uses of some species' leaves, shade provision and hedging, as well as to the fact that farmers who have more tropical livestock units are also found to be more depending on farming as an income source. The positive effects of natural tree occurrence and amount of parcels are already explained in Section 2.4.1, where they were proposed as selection variables.

In two-stage least square estimations with "Group membership" and "Media exposure" as instruments, both proved to be relevant ($\beta_{1st\ stage}$: 0.129 ($p = 0.003$) and 0.028 ($p = 0.084$) respectively) and their validity (exogeneity) was not rejected (p-value of overidentification test: 0.440). However, including "Media exposure" makes the instrument set weaker; therefore, the second stage control function estimates in Table 5 are based on a first stage using only "Group membership" as an instrument ($\beta_{1st\ stage} = 0.135$ ($p = 0.003$)). In this second stage, the estimate of the main BP training effect is similar to the base model, but due to a large standard error it is not significant anymore. Therefore, although the effect of BP training is still significantly different for farmers who did not have trees in 2007 compared to those who did, for both it is not significant anymore. However, the coefficient of the first stage residuals is insignificant, or alternatively, the null hypothesis that there is no self-selection bias, is not rejected[13]. This is also reflected by the estimates for the control variables, which are almost identical to the base model. It indicates that the estimates of the base model are valid. Lastly, and as in the base model, p-values based on heteroscedasticity-robust standard errors and p-values based on wild cluster bootstrapped t-statistics indicate the same variables to be significant, and only differ in terms of the corresponding significance levels.

3.3.2. Seed collection

The seed collection equation aims primarily to quantify the direct effect of BP training on adoption. However, there exists an indirect effect through tree cultivation: BP training has a positive effect on tree cultivation (see Section 3.3.1), and tree maturity very likely has a positive effect on seed collection (see Section 3.1). To estimate this indirect effect, tree maturity is added as a regressor to the equation. Furthermore, it is also interacted with BP training. The reasoning is the following. One can reasonably expect that many farmers not having mature trees simply do not have the ability (or permission) to collect seeds within a reasonable distance (only 2.4% collects, see Section 3.1), even if they would be willing to. BP training is therefore simply unable to make a difference for these households. The interaction will prevent this mechanism from "diluting" the real (or potential) effect of BP training. The linear probability model estimates for this specification, using village fixed effects, are presented in Table 6. As hypothesized, the base model estimates reveal a significant positive effect of tree maturity, while the effect of BP training is also higher for farmers having mature trees compared to farmers who do not (for regular p-values). However, for both groups, the overall BP training effect is not found to be significant (main + interaction effect: 0.14 ($p = 0.122$) ; main effect: -0.04).

Next, among the control variables only the assets index is found to be a significant determinant of adoption. It indicates that wealthier households are less likely to get involved into seed collection activities. This corresponds to the traditional notion of oilseed collection as a marginal activity.

[13] This is also the case in the two-stage least squares based test for endogeneity.

More technically, it may indicate that for wealthier households the opportunity costs for land, capital and especially labour are too high to get involved into seed collection, although explicit proxies for these costs are also present in the model (number of HH labourers, exploited land) but not found to be significant. In the control function model, natural tree occurrence is additionally found to have a negative effect (for bootstrapped p-values), as in Section 3.2. This may seem strange at first sight, but it should be re-emphasized that tree presence is already controlled for through "Mature trees". Rather, it refers to farmers who maintain these wild trees probably for their multiples uses and co-benefits, but not for seed collection. This again emphasizes the need to explicitly distinguish between tree cultivation and seed collection.

The results of the instrument-based analysis are very similar to those in the tree cultivation model (Section 3.3.1). "Group membership" and "Media exposure" are relevant instruments and their validity (exogeneity) is not rejected, but the latter is a weak instrument and therefore not used for the second stage control function estimates in Table 6. In this second stage, the standard error of the main BP training effect is again largely inflated. The BP training effect is now also negative for farmers who have mature trees (main + interaction effect: -0.11 ($p = 0.705$)) as for those who do not (main effect: -0.29), but insignificant. However, due to the insignificance of the first stage residuals, there is again no indication of self-selection bias[14], which indicates that the estimates of the base model are valid.

3.4. The effect of cooperatives and expellers

The base models of Section 3.3 were also estimated with hobli instead of village fixed effects, while adding "BP association" and "BP expeller" as regressors (data not shown). However, even when the simultaneity bias introduced by these predictors is acknowledged, the bias direction is complicated due to regressor correlations. Therefore, only the estimates for these predictors themselves are assessed, with the simplifying assumption that they are not correlated with other regressors (Wooldridge, 2010). In that case, a positive bias would be expected, especially for BP expeller. Yet it does not have significant effects on tree cultivation and seed collection adoption probabilities, while BP association even has a significant negative effect on seed adoption probabilities, in line with the descriptive statistics (Table 2).

[14] This is also the case in the two-stage least squares based test for endogeneity.

Table 5. Estimates for tree cultivation adoption (y_t): base model using ordinary least squares (OLS) and 2nd stage of the control function model using group membership as instrument (CF).

	OLS				CF - 2nd stage			
	β	(SE)	p	p-btrap	β	(SE)	p	p-btrap
Gender of HH head (1 = male)	-0.017	(0.059)	0.773	0.772	-0.017	(0.058)	0.775	0.771
Age of HH head (years)	0.000	(0.002)	0.979	0.979	0.000	(0.002)	0.976	0.976
Education of HH head	**0.114**	**(0.040)**	**0.005**	**0.009**	**0.114**	**(0.040)**	**0.005**	**0.011**
Number of HH dependents	0.004	(0.026)	0.882	0.887	0.005	(0.030)	0.879	0.885
(Number of HH dependents)²	0.002	(0.004)	0.599	0.599	0.002	(0.005)	0.710	0.724
Number of HH labourers	0.108	(0.066)	0.104	0.136	0.109	(0.066)	0.101	0.136
(Number of HH labourers)²	-0.016	(0.010)	0.115	0.192	-0.016	(0.010)	0.114	0.194
Exploited land (10^{-1} acres)	0.046	(0.104)	0.661	0.609	0.045	(0.105)	0.666	0.620
(Exploited land)² (10^{-1} acres)^2	-0.022	(0.029)	0.450	0.392	-0.021	(0.032)	0.505	0.515
Amount of tropical livestock units	**0.070**	**(0.024)**	**0.004**	**0.003**	**0.070**	**(0.027)**	**0.011**	**0.015**
(Amount of tropical livestock units)²	**-0.008**	**(0.003)**	**0.015**	**0.087**	**-0.008**	**(0.004)**	**0.063**	**0.094**
Assets and living standards index	-0.014	(0.009)	0.115	0.118	-0.014	(0.009)	0.134	0.164
Natural tree occurence (1 = yes)	**0.286**	**(0.047)**	**0.000**	**0.000**	**0.286**	**(0.051)**	**0.000**	**0.000**
Amount of parcels	**0.037**	**(0.015)**	**0.015**	**0.023**	**0.038**	**(0.019)**	**0.051**	**0.061**
Trees in 2007 (1 = yes)	**0.403**	**(0.052)**	**0.000**	**0.000**	**0.404**	**(0.055)**	**0.000**	**0.000**
BP training (1 = yes)	**0.228**	**(0.087)**	**0.009**	**0.099**	0.213	(0.347)	0.540	0.624
Trees in 2007*BP training	**-0.322**	**(0.103)**	**0.002**	**0.036**	**-0.322**	**(0.104)**	**0.002**	**0.038**
Constant	0.004	(0.151)	0.978	0.976	0.002	(0.152)	0.990	0.990
1st stage residuals					0.015	(0.344)	0.964	0.972
Village fixed effects	(included)				(included)			
R²	0.631				0.631			
Amount of observations	390				390			

Table 6. Estimates for seed collection adoption (y_s): base model using ordinary least squares (OLS) and 2nd stage of the control function model using group membership as instrument (CF).

	OLS				CF - 2nd stage			
	β	(SE)	p	p-btrap	β	(SE)	p	p-btrap
Gender of HH head (1 = male)	-0.016	(0.054)	0.765	0.747	-0.012	(0.055)	0.825	0.813
Age of HH head (years)	0.000	(0.002)	0.936	0.927	0.000	(0.002)	0.986	0.984
Education of HH head	-0.049	(0.041)	0.237	0.141	-0.047	(0.042)	0.263	0.165
Number of HH dependents	-0.022	(0.026)	0.402	0.441	-0.010	(0.028)	0.725	0.707
(Number of HH dependents)²	0.003	(0.004)	0.501	0.567	0.001	(0.005)	0.905	0.902
Number of HH labourers	-0.075	(0.063)	0.233	0.250	-0.069	(0.063)	0.272	0.287
(Number of HH labourers)²	0.010	(0.009)	0.302	0.278	0.009	(0.009)	0.312	0.289
Exploited land (10^{-1} acres)	-0.048	(0.098)	0.623	0.663	-0.057	(0.099)	0.564	0.600
(Exploited land)² (10^{-1} acres)^2	-0.011	(0.029)	0.709	0.732	-0.002	(0.030)	0.956	0.956
Amount of tropical livestock units	0.017	(0.018)	0.351	0.365	0.024	(0.020)	0.245	0.243
(Amount of tropical livestock units)²	-0.002	(0.002)	0.271	0.228	-0.004	(0.003)	0.178	0.140
Assets and living standards index	**-0.022**	**(0.009)**	**0.014**	**0.010**	**-0.020**	**(0.010)**	**0.042**	**0.023**
Natural tree occurence (1 = yes)	-0.044	(0.036)	0.218	0.143	**-0.059**	(0.040)	0.136	**0.081**
Amount of parcels	0.009	(0.017)	0.586	0.699	0.018	(0.019)	0.337	0.477
Mature trees (1 = yes)	**0.287**	**(0.053)**	**0.000**	**0.000**	**0.309**	**(0.057)**	**0.000**	**0.000**
BP training (1 = yes)	-0.035	(0.039)	0.362	0.210	-0.288	(0.304)	0.344	0.239
Mature trees*BP training	**0.174**	**(0.095)**	**0.068**	0.157	**0.174**	**(0.095)**	**0.068**	0.157
Constant	0.181	(0.157)	0.249	0.250	0.146	(0.158)	0.355	0.343
1st stage residuals					0.259	(0.304)	0.393	0.298
Village fixed effects	(included)				(included)			
R^2	0.338				0.339			
Amount of observations	390				390			

3.5.Reflections on adoption rates and the effect of the Biofuel Program

The primary aim of the study is to investigate tree cultivation and seed collection adoption rates, and how this is affected by BP extension and implementation. While the majority of farmers cultivate oilseed trees, seed collection rates are generally low. BP extension ("BP training") also only stimulates farmers to cultivate trees. Complementing these observations with additional qualitative information from the survey, two fundamental reasons for these findings are proposed.

First, adoption of these practices involves opportunity costs of land, labour and capital (see also Dalemans et al. (in press) and van Eijck et al. (2014)). Even though the agroforestry approach targets underutilized lands and involves minimal management practices, respondents still mention that these systems are not profitable, not useful, and interfere with crop production, while also stating that they lack labour, land and/or water. Costs may still be limited for tree cultivation, especially if extension is supplemented with free seedling distribution programs, while respondents also value the co-benefits (e.g., hedging, shading) and multiple uses (e.g., compost, furniture) in an agroforestry set-up. Seed collection however requires extensive labour, and while its role could not explicitly be identified in the econometric model (see the discussion on the asset index in Section 3.3.2), lack of labour is often explicitly mentioned as a barrier to adoption, or indirectly as "seed collection is not profitable". The latter indicates that willingness to collect seeds is highly depending on the value chain. As respondents also mention that a lack of markets and market information form a barrier to adoption, the value chain is likely still underdeveloped.

This also relates to the second reason: BP extension may be not intensive enough to stimulate adoption (see also Basinger et al. (2012) and Glendinning et al. (2001)). In an effort to cover as many villages as possible with limited resources, the BP has only implemented one or few trainings in many villages, and these might have taken place years ago. This could explain the mentioned lack of both information and marketing possibilities. Furthermore, it could explain why no effects of cooperatives and expeller distributions were found. More specifically, the striking observation that none of the households in cooperative villages are aware about the biodiesel cooperatives, is probably because they are simply not operational (anymore). Similarly, the fact that only one household in the sample uses an oil expeller, and only one household sells oilseeds to the BP, probably results from a lack of on-site support and promotion. Processing rates of these small-scale expellers are also very low, which again compromises profitability.

4. Conclusions and Policy implications

This study addresses the lack of empirical evidence on farmer adoption of biofuel systems. For a case study in Hassan district, India, farmer perception and adoption rates are described, while the roles of a biofuel extension program (BP) and farm(er) characteristics are quantified through various econometric models.

The findings reveal that while 60% of the farmers cultivate oilseed trees, a large discrepancy exists with oilseed collection (78% of tree cultivators do not collect seeds) – even when accounting for tree immaturity (69%) – and that adoption of these two practices is also driven by different determinants. They indicate that adoption of tree cultivation and seed collection should be considered as two related stages rather than a single binary choice. The latter would neglect several inherent mechanisms and characteristics of the adoption process. Consequently, results could be biased and little informative for fundamental understanding and policy recommendations.

Not only is oilseed tree cultivation much more prevalent than oilseed collection, BP awareness programs and trainings are also found to only stimulate the former and not the latter. Therefore, the BP seems to be succeeding as an agroforestry program, which is likely driven by the limited (opportunity) costs, free seedling distribution programs, and farmers' valuation of the multiple uses and co-benefits of oilseed trees. However, the BP does not seem to be succeeding as a biofuel program. This is also confirmed by the inability of advanced program implementations, such as biodiesel cooperative formation and oilseed expeller distribution, to increase seed collection and local processing rates. The latter is only done by one sampled household, while almost all others market the seeds in traditional, non-biofuel value chains. These observations are likely driven by high opportunity costs for land, water and especially labour, an underdeveloped value chain, and lack of BP implementation and on-site support.

Although the BP may increase seed collection rates in the long term indirectly through increased tree cultivation (and maturity) rates, the lack of a direct effect as well as the almost complete lack of seed sourcing for biofuel production demand the BP to be revised. First, the BP should acknowledge that even in small-scale, low-input systems, opportunity costs of land, labour and capital play a crucial role (see also Dalemans et al., in press; van Eijck et al., 2014). This emphasizes the need for a profitability analysis, which could in turn imply the need to strengthen value chains, modify business models, differentiate program targeting and/or dissolve the program (or its biofuel component). Second, but only conditional on these results, the BP should intensify extension and implementation. Given limited resources, they could consider to refrain from a strategy of reaching as many villages as possible, and rather intensify their activities in clusters of villages, from which it may diffuse passively.

5. References

Achten, W.M.J. et al., 2014. Opportunities and constraints of promoting new tree crops - lessons learned from jatropha. Sustainability 6, 3213–3231.

Altenburg, T. et al., 2009. Biodiesel in India: value chain organisation and policy options for rural development. Studies 43, German Development Institute, Bonn, Germany.

Axelsson, L. et al., 2012. Jatropha cultivation in southern India: assessing farmers' experiences. Biofuels, Bioprod. Biorefining 6, 246–256.

Basinger, M. et al., 2012. Jatropha adoption: A statistical observational study of factors influencing Malian farmers' decision to grow Jatropha. Agrofor. Syst. 84, 59–72.

Biswas, P.K., Pohit, S., 2013. What ails India's biodiesel programme? Energy Policy 52, 789–796.

Cameron, A.C., Miller, D.L., 2015. A practitioner's guide to cluster-robust inference. J. Hum. Resour. 50, 317–372.

DAC&FW, 2017. Agriculture census 2010-2011. Retrieved from http://agcensus.nic.in.

Dalemans, F. et al., in press. Redesigning oilseed tree biofuel systems in India. Energy Policy.

DCO, 2014. District census handbook Hassan. Census of India 2011. Directorate of Census Operations, Bangalore, India.

Demirbas, A.H., Demirbas, I., 2007. Importance of rural bioenergy for developing countries. Energy Convers. Manag. 48, 2386–2398.

Ewing, M., Msangi, S., 2009. Biofuels production in developing countries: assessing tradeoffs in welfare and food security. Environ. Sci. Policy 12, 520–528.

FAO et al., 2017. The State of Food Security and Nutrition in the World 2017. Building resilience for peace and food security. Rome.

Feder, G. et al., 1985. Adoption of Agricultural Innovations in Developing Countries: A Survey. Econ. Dev. Cult. Change 33, 255.

Glendinning, A. et al., 2001. Modes of communication and effectiveness of agroforestry extension in Eastern India. Hum. Ecol. 29, 283–305.

Greene, W.H., 2012. Econometric analysis. Pearson.

Heckman, J.J., 1976. The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. Ann. Econ. Soc. Meas. 5, 475–492.

IEA, 2015. India Energy Outlook, World Energy Outlook Special Report. International Energy Agency, Paris, France.

Imbens, G., Wooldridge, J., 2007. What's new in econometrics.

Mercer, D.E., 2004. Adoption of agroforestry innovations in the tropics: A review. Agrofor. Syst. 61–62, 311–328.

Pattanayak, S.K., et al., 2003. Taking stock of agroforestry adoption studies. Agrofor. Syst. 57, 173–186.

Sharma, N. et al., 2016. Bioenergy from agroforestry can lead to improved food security, climate change, soil quality, and rural development. Food Energy Secur. 5, 165–183.

Sileshi, G. et al., 2007. Contributions of agroforestry to ecosystem services in the Miombo eco-region of eastern and southern Africa. African J. Environ. Sci. Technol. 1, 68–80.

Soto, I. et al., 2015. Who benefits from energy policy incentives? The case of jatropha adoption by smallholders in Mexico. Energy Policy 79, 37–47.

Tilman, D. et al., 2006. Carbon-negative biofuels from low-input high-diversity grassland biomass. Science (80-.). 314, 1598–1600.

Van de Ven, W.P.M.M., Van Praag, B.M.S., 1981. The demand for deductibles in private health insurance. J. Econom. 17, 229–252.

van Eijck, J. et al., 2014. Global experience with jatropha cultivation for bioenergy: an assessment of socio-economic and environmental aspects. Renew. Sustain. Energy Rev. 32, 869–889.

Wooldridge, J.M., 2015. Control Function Methods in Applied Econometrics. J. Hum. Resour. 50, 420–445.

Wooldridge, J.M., 2014. Quasi-maximum likelihood estimation and testing for nonlinear models with endogenous explanatory variables. J. Econom. 182, 226–234.

Wooldridge, J.M., 2010. Econometric analysis of cross section and panel data. MIT press.