

RESEARCH

Open Access



# Impact of dairy hubs on smallholder welfare: empirical evidence from Kenya

Vincent Ngeno

Correspondence: [vcngeno@gmail.com](mailto:vcngeno@gmail.com)

Department of Agricultural  
Economics & Resource  
Management, Moi University,  
Eldoret, Kenya

## Abstract

This paper identifies the factors that affect farmers' decision to participate in a dairy hub. It also examines the impacts of dairy hub participation on rural household welfare measured by farm yields and net returns in Kenya. The study utilises farm-level data of 826 Kenyan small-scale dairy farmers. The casual impact of dairy hub participation is estimated by utilising endogenous switching regression. This helps in estimating the true welfare effect of dairy hub participation by controlling the role of selection problem on participation decision. The empirical results indicate that farmers' education, extension services, credit access, landholding, tropical livestock unit, distance to dairy hub, off-farm work and household size mainly determine participation in dairy hub. It was also found that dairy hub participation significantly increases yields and net returns.

**Keywords:** Household welfare, Dairy hub participation, Impact assessment, Endogenous switching

## Background

Smallholders, owning one to three cows, account for about 80% of the milk production in Kenya (Makoni et al. 2013). Livestock is, therefore, an important asset for many Kenyan rural residents. Kenya is experiencing a growing demand for milk and dairy products driven by their strong tradition of including milk in diets, population growth, expanding urbanisation, a rising middle class and export opportunities (Rademaker et al. 2016). The potential of the Kenyan dairy sector to increase production to meet consumer demands of the growing population and enhance livelihood benefits has not gone unnoticed (Vernooij 2016). Although the growing demand for milk products in Kenya is widely recognised, the dairy sector has failed to produce adequate milk to satisfy this demand, mainly due to challenges in dairy farming.

These challenges are largely similar across other developing countries and include, among others, the following problems: (1) low productivity driven by limited access to quality and affordable inputs and services, and output markets (Rao et al. 2016b); (2) limited access to input market that heightens the cost of production, further restricting households to low-input-low-output vicious cycle (Rao et al. 2015); (3) remote and scattered location of smallholder dairy households with limited access to reliable infrastructure, which leads to higher transaction costs and presents further challenges in terms of access to input and output markets (Rao et al. 2016a); (4) inappropriate cattle

breeds for dairy production; (5) diseases and low adoption of technologies in the dairy industry (Makoni et al. 2013); and (6) high transaction cost (Kilelu et al. 2016).

The dairy hub model is an innovative approach developed to address some of these challenges. Dairy hub is a collective farmer-owned/managed milk bulking and/or chilling business from which farmers may also gain access to other services they need for their daily enterprises (Mutinda et al. 2015). The East Africa Dairy Development project, run by Heifer International and other development partners, has successfully rolled out dairy hubs in Kenya. One of the project objectives is to boost the milk yields and net returns of small-scale farmers in Africa (Kenya, Uganda and Tanzania) (Pyburn and Eerdewijk 2016). Dairy hubs have proven to be potentially strong platforms for improving access to markets, inputs and services for both men and women smallholder dairy farmers alike (Rademaker et al. 2016). Indeed, they are transforming rural regions (Kilelu et al. 2016; Mutinda et al. 2015).

Although dairy hubs offer a number of benefits, the rate of participation in these dairy hubs has been slow so far (Omondi et al. 2016). There is limited empirical evidence on the determinants of participation of smallholder dairy farmer in dairy hubs or the effects of this participation on smallholder welfare. A recent study in East Africa by Rao et al. (2016a, b) employed propensity score matching (PSM) for examining the impacts of different marketing strategies adopted by dairy business hubs on dairy and household income. However, a well-known shortcoming of PSM method is its inability to account for unobservable factors, resulting in biased estimates. Muchichu (2014) assessed only the determinants of the adoption of dairy hubs among smallholder farmers in rural Tanzania without considering its impacts, yields or net returns.

There is a lack of empirical studies that have evaluated the impact of dairy hubs on smallholder welfare using endogenous switching regression (ESR). The ESR approach accounts for selection bias due to both observed and unobserved heterogeneity. ESR allows the distribution of the endogenous treatment (participation in dairy hubs) and outcomes (net returns and yields) to be specified using a latent factor structure, thereby differentiating between selection on unobservables and selection on observables (Deb and Trivedi 2006). This study aims to bridge the research gap by applying ESR to identify the determinants that influence smallholders' participation in dairy hubs, as well as investigate the effects of this participation on their welfare.

### **Data description**

The data employed in this study were collected from dairy farming households in hubs' catchment areas in the highlands of Rift Valley in Kenya. A multistage sampling procedure was used for the selection of observation units. First, the catchment areas serving the Siongiroi dairy hub in South Rift and Kibiyet dairy hub in North Rift were purposely selected. In these two catchment areas, seven administrative locations in each catchment area were purposely selected using the statistical information on milk production. Finally, 59 dairy farmers were randomly selected in each location. In total, the data set comprises observation from 826 farm households—457 that participated in dairy hubs and 369 that sold in traditional channels.

A structured household questionnaire was used for data collection. The questionnaire gathered information on a range of relevant topics such as household

characteristics, details on milk production and marketing, as well as social and institutional characteristics.

Table 1 presents the definition of selected variables and the mean difference between participants and non-participants in dairy hubs. It reveals that there are significant differences between participants and non-participants with respect to several characteristics. This is because farmers self-select into the group of dairy hub participants according to their conditions and preferences, which needs to be accounted for in the impact analysis. The dependent variable used in the study is a dummy variable that takes the value of one if the household participated in dairy hub and the value zero if no participation was recorded. The outcome variables used in this study are milk yields and net returns. With regard to the outcome variables, there appear to be statistically significant differences in milk yields and net returns between participants and non-participants. As evident from Table 1, the average milk yield of participants is higher than that of non-participants. Moreover, the net returns are significantly higher for participants. It implies that dairy hubs might have contributed to net returns.

In particular, statistically significant differences between participants and non-participants include age and education of the household head and household's

**Table 1** Descriptive statistics of selected variables for participants and non-participants

Variable	Description of variable	Mean		Difference in means
		Participants	Non-participants	
Treatment variable				
Participant	1 if farmer participates in dairy hub, 0 otherwise			
Outcome variables				
Milk yields	Milk output per cow/day/litre	4.93	3.90	1.03**
Net returns	Milk gross revenue minus variable investment costs (1000 Ksh/year)	473	317	156**
Independent variables				
Age	Age of household head (years)	46.4	49.1	-2.7***
Education	Years of schooling of household head	8.7	7.6	1.1***
Experience	Dairy farming experience (years)	19.4	23.6	-4.2**
Ext. service	1 if farmer has contact with extension agent, 0 otherwise	0.613	0.431	0.182
Credit access	1 if farmer has access to credit, 0 otherwise	0.37	0.17	0.20***
Landholding	Size of landholding in hectare	2.9	1.8	1.1**
Livestock	Total cattle owned (tropical livestock unit)	6.87	3.79	3.08
Distance	Distance to hub (km)	7	13	-6***
Sex	1 if male-headed, 0 otherwise	0.818	0.661	0.157***
Off-farm	1 if the household participated in off-farm employment, 0 otherwise	0.466	0.416	0.050
HH size	Number of adult household members in adult	6	4	2***
Dependents	Number of household aged < 14 and > 64 years	3	4	-1
Fgroup	Membership in farmers' group	67.58	47.39	20.19***
DistExt	Distance to extension office (minutes of walking time)	62	102	-40

Exchange rate US \$1 = 100 Ksh in 2017

Notes: Ksh, Kenyan Shillings

\*\*\*Mean difference significant at 1%

\*\*Mean difference significant at 5%

experience in milk production. The dairy hub participants are younger but have more a formal education while the non-participants have more experience in producing milk, perhaps because they tend to be older. There is a statically significant difference in access to credit between participants and non-participants. Thirty-seven percent of participants and 17% of non-participants had access to formal credit, revealing the low levels of credit availability to rural farm households.

Landholding is an indicator of the level of household resource endowment, with larger landholding indicating more household resource endowment. The average landholding of the participants is significantly higher compared to the non-participants. Distances to the market are significantly lesser for the participants compared to the non-participants. The gender composition of participants is significantly more male-dominated than that in non-participants. On average, participants have a significantly larger household size than non-participants. With respect to dependency, participants supported a fewer number of people who are either young or old compared to non-participants. There appears to be a significant difference between participants and non-participants in terms of membership in farmers' organisations. Overall, the results in Table 1 indicate that participants and non-participants are systematically different.

**Methods**

Estimation of the differences in the welfare of farm households of the participants and non-participants based on non-experimental observation is not trivial because of the need to identify the counterfactual situation had they not participated in dairy hub. In experimental studies, this problem is addressed by randomly assigning farmers to treatment and control status, where the outcome observed on the control households (non-participants) are statistically representative of what would have occurred without participation for treated farmers (participants). However, farmers are not randomly distributed to the two groups (participants and non-participants), but rather farmers make their own adoption choices or are systematically selected by development agencies and/or project administrators based on their propensity to participate in dairy hubs. Therefore, participants and non-participants may be systematically different. Thus, possible self-selection due to observed and unobserved plot and household characteristics make it difficult to perform ex-post assessment of gains from dairy hub participation using observational data.

To address the above econometric challenges, the ESR method was proposed. ESR is a two-stage framework used to model household decisions to participate in dairy hubs (first stage) and the impact of participation on the welfare of farm households (second stage).

In the first stage, a probit model is applied to estimate the determinants of participation in dairy hubs. A farm household *i* chooses to participate in a dairy hub if it generates the perceived net benefit. This is specified as follows (Mojo et al. 2017; Wossen et al. 2017):

$$\delta_i^* = \gamma X_i + \varepsilon_i \text{ with } \delta_i^* = \begin{cases} 1 & \text{if } \delta_i^* > 0 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where  $\delta_i^*$  is the latent variable that captures the expected benefit from participating in dairy hubs.  $\gamma$  is a vector of parameters to be estimated,  $X_i$  is a vector of observed covariates and  $\varepsilon_i$  is an error term assumed to be normally distributed with zero mean. The observed dependent variable, i.e., participation status, is denoted as  $\delta_i$ , where  $\delta_i = 1$  for

participants and  $\delta_i = 0$  for non-participants (it is also denoted as  $\delta_i^*$ ) as indicated in Eq. (1). The probability of being a participant of a dairy hub can be expressed as:

$$\Pr(\delta_i = 1) = \Pr(\delta_i^* > 0) = \Pr(\varepsilon_i > -\gamma X_i) = 1 - F(-\gamma X_i) \tag{2}$$

where  $F$  is the cumulative distribution function for  $\varepsilon_i$ .

In the first stage, the decision of whether or not to participate in dairy hubs is estimated. From Eq. (1), which is also referred to as the selection equation, the inverse Mills ratio is computed and included in the second stage of the equations to control for selectivity bias.

In the second stage, the differential impact of covariate on the welfare of farm households measured by milk yields and net returns of the group of participants and non-participants is modelled simultaneously through an endogenous switching regime model. Conditional on the dairy hub participation decision being observed in the first stage, the welfare of farm households in the second stage is estimated as two separate regimes as specified below (Fuglie and Bosch 1995; Tambo and Wünsch 2017):

$$\begin{aligned} y_1 &= \vartheta_1 X + \sigma_{\mu_1, \varepsilon} \lambda_1 + \xi_1 & \text{if } \delta = 1 \\ y_0 &= \vartheta_0 X + \sigma_{\mu_0, \varepsilon} \lambda_0 + \xi_0 & \text{if } \delta = 0 \end{aligned} \tag{3}$$

where  $y_1$  and  $y_0$  represent a vector of welfare of farm households for dairy hub participants and non-participants, respectively.  $\vartheta_1$  and  $\vartheta_0$  are parameters to be estimated for participants and non-participants regimes, respectively.  $X$  is a vector of explanatory variables that influence the welfare of farm households.  $\lambda_1$  and  $\lambda_0$  are the inverse mill ratios evaluated at  $\gamma X$ ;  $\sigma_{\mu_1, \varepsilon} = \text{cov}(\mu_1, \varepsilon)$  and  $\sigma_{\mu_0, \varepsilon} = \text{cov}(\mu_0, \varepsilon)$ . When the error term of the selection equation is correlated with the error terms of the outcome equation of participants and non-participants, then we have a selection bias problem. Thus, estimates from the selection equation are used to compute  $\lambda_1$  and  $\lambda_0$ , which are then added to the outcome equations to correct for selection bias. This can be estimated using a two-stage method (Maddala 1986). However, the full information maximum likelihood (FIML) estimation approach (Lokshin and Sajaia 2004) is used as it estimates the selection and outcome equations simultaneously. This is more efficient than the two-step procedure. If  $\sigma_{\mu_1, \varepsilon}$  and  $\sigma_{\mu_0, \varepsilon}$  in Eq. (3) are statistically significant, endogenous switching exists.

While the FIML ESR model is identified through non-linearities of  $\lambda_1$  and  $\lambda_0$  (Lokshin and Sajaia 2004), a better identification requires an exclusion restriction. In other words, for the ESR model to be correctly specified, the selection equation should contain at least one selection instrument in addition to those generated by the non-linearity of the selection model correlated with dairy hub participation but uncorrelated directly with milk yields and net returns. Following the literature (see Khonje et al. 2015a, b), the selection instrument used is distance to agriculture extension office. The validity of the instrument is tested using falsification tests following previous studies. The results show that the selected instruments can be considered as valid as it is statistically significant in explaining participation decision [ $\chi^2 = 8.71$  ( $p = 0.000$ )] but is not statistically significant in explaining the milk yield function [ $F = 1.71$  ( $p = 0.514$ )] and [ $F = 1.07$  ( $p = 0.214$ )] for participants and non-participants, respectively. Taking net returns as an outcome variable, the falsification test also shows an insignificant effect for participants [ $F = 2.27$  ( $p = 0.454$ )] and non-participants [ $F = 2.87$  ( $p = 0.116$ )], verifying the validity of the instrument.

The coefficients from the ESR model can be used to derive the expected values of welfare of farm households, which are used in estimating the average treatment effect on the treated (ATT):

$$ATT = E(y_1|\delta = 1) - E(y_0|\delta = 1) = X(\vartheta_1 - \vartheta_0) + \lambda_1(\sigma_{\mu_1\varepsilon} - \sigma_{\mu_0\varepsilon}) \quad (4)$$

The ATT compares the welfare of farm household indicators of dairy hub participants with and without dairy hubs.

## Results and discussion

The descriptive statistics in the previous section revealed significant differences in selected variables between participants and non-participants in dairy hubs. Yet to properly analyse impacts, econometric approaches are needed. As outlined in the “[Empirical specification](#)” section, an ESR is applied to estimate the effects of yield and net returns on participation in dairy hubs. The FIML approach is employed to estimate the outcome equations jointly with the selection equation that explains farmers’ participation in dairy hubs.

The estimates of the determinants of participation and the impact of participation on milk yields and net returns are presented in Tables 2 and 3, providing two different sets of results due to slightly different specifications. As indicated previously, the FIML is employed to jointly estimate the selection and outcome equations. Thus, columns (2) and (3) present the estimated coefficients and standard errors of the selection equation, representing determinants of participating in dairy hubs. The outcome equations are presented in columns (4) to (6). In the following, the results on the determinants of participation are first discussed before focusing on yield and net returns effects. The group membership estimates are not statistically significant.

### Determinants of dairy hub participation

Given that the empirical results in the selection equation can be interpreted as normal probit coefficients, the results for variables with the same name from the selection equation presented in Tables 2 and 3 have similar effects on participation. The variable representing education of the farmer is positive and significantly different from zero, suggesting that more educated farmers are more likely to participate in dairy hubs. This finding is consistent with the notion that farmers with more education are more likely to be more confident in adjusting to new innovations and technologies (Abdulai and Huffman 2014; Rao and Qaim 2011).

Results show that access to extension services increases the likelihood of participation. Farmers who are regularly visited by extension workers and those who attend field days or host demonstrations/trials or get media extension messages are likely to participate due to their increased exposure and awareness (Khonje et al. 2015a). This result confirms the importance of information as a means of reducing uncertainty about innovations. In Kenya, agricultural extension tends to be a major source of information on technological improvement and can be crucial to the participation in dairy hubs.

The access to credit variable, an indicator of liquidity constraint, is positive and significantly different from zero, suggesting that liquidity-constrained farmers are less likely to participate. This confirms the significance of access to credit in participation decision process. Access to credit for various inputs enables farmers to easily

**Table 2** Determinants of dairy hub participation and the impact of participation on milk yields

Variable	Selection		Milk yields			
			Participants		Non-participants	
	Coefficient	Std. error	Coefficient	Std. error	Coefficient	Std. error
Constant	0.387	0.524	1.254**	0.531	-1.682	3.571
Age	-0.161	0.109	-0.216	0.378	0.076	0.056
Education	0.439***	0.127	3.374**	1.282	2.019***	1.015
Experience	0.237	0.145	0.296	0.182	0.624	0.403
Ext. service	0.459**	0.206	2.768**	1.123	5.520**	2.361
Credit access	0.071**	0.03	0.056**	0.022	0.039***	0.014
Landholding	0.178**	0.07	0.735	0.498	-0.627*	0.331
Livestock	0.199**	0.101	0.047***	0.016	0.593	0.376
Distance	-0.171***	0.04	-0.076***	0.02	-0.088***	0.01
Sex	0.035	0.27	-0.237	0.16	-0.058	0.13
Off-farm	-0.256**	0.11	0.104	0.07	0.091***	0.03
HH size	-0.209***	0.08	0.179**	0.09	-0.217	0.15
Dependents	-0.172	0.11	-0.336	0.683	0.187	0.127
F group	0.008	0.037	0.085	0.089	0.071	0.061
Res(DistExt)	0.073	0.061			0.427***	0.119
$\ln\sigma_{NP}$						
$\rho_{NP}$					-0.177	-1.492
$\ln\sigma_P$			0.627***	0.143		
$\rho_P$			-0.537***	0.184		
Log likelihood	-485.438					
$\chi^2$ statistics for overidentification				0.605 [0.41]		
Likelihood ratio test of independent equations $\chi^2(1)$				17.63***		

Note: *p* value in square brackets. Res(DistExt)—RES denote residuals from the first-stage regressions for distance to agriculture extension office

\*Significance at the 10% level

\*\*Significance at the 5% level

\*\*\*Significance at the 1% level

participate in dairy hub unlike in cases where it is a constraint. Therefore, increased access to institutional support services such as extension, credit, and input supply should be a major part of efforts aimed at promoting dairy hub participation, as pointed out by Khonje et al. (2015a, b). Land holding has a significant positive effect on the likelihood of dairy hub participation. Larger farms are at a certain advantage, which may be due to fixed transaction costs in dairy hub channel such as transportation costs and information search costs (Rao and Qaim 2011).

The results show a positive and significant relationship between tropical livestock units (TLU) and dairy hub participation. For participants, livestock enables them to have additional resources that can be used in financing farming resources. Livestock size is often used as an indicator of wealth in agrarian settings. Our results suggest that wealthier households are more likely to participate in dairy hubs because they may be able and willing to bear more risks than their counterparts and may have preferential access to inputs and credit (Kassie et al. 2009).

The distance to dairy hub negatively and significantly influences participation in dairy hub. As travel time and transportation costs increase, farmers are more likely to sell their milk through a traditional channel. Off-farm income is negatively related to dairy

**Table 3** Determinants of dairy hub participation and the impact of participation on net returns

Variable	Selection		Net returns			
			Participants		Non-participants	
	Coefficient	Std. error	Coefficient	Std. error	Coefficient	Std. error
Constant	0.288	0.411	1.365**	0.640	-1.573	2.863
Age	-0.252	0.167	-0.205	0.467	0.094	0.073
Education	0.394**	0.201	3.437***	1.196	1.901**	0.971
Experience	0.372	0.415	0.269	0.167	0.815	0.587
Ext. service	0.395**	0.167	2.876**	1.132	4.215**	1.713
Credit access	0.127**	0.057	0.065**	0.027	0.127**	0.063
Landholding	0.187**	0.087	0.537	0.378	-0.739*	0.413
Livestock	0.197**	0.093	0.107**	0.049	0.359	0.265
Distance	-0.243***	0.089	-0.144***	0.039	-0.179***	0.067
Sex	0.054	0.73	-0.324	0.197	-0.036	0.137
Off-farm	-0.345**	0.17	0.213	0.173	0.113***	0.037
HH size	-0.119***	0.043	0.239**	0.109	-0.172	0.151
Dependents	-0.261	0.173	-0.248	0.551	0.278	0.169
F group	0.157	0.149	0.329	0.231	0.158	0.126
Res(DistExt)	0.134	0.127				
$\ln\sigma_{NP}$					0.359***	0.104
$\rho_{NP}$					-0.375	-1.294
$\ln\sigma_P$			0.818***	0.241		
$\rho_P$			-0.459***	0.148		
Log likelihood	-584.348					
$\chi^2$ Statistics for overidentification				0.565 [0.41]		
Likelihood ratio test of independent equations $\chi^2(1)$				27.37***		

Note: *p* value in square brackets. Res(DistExt)—Res denote residuals from the first-stage regressions for distance to agriculture extension office

\*Significance at the 10% level

\*\*Significance at the 5% level

\*\*\*Significance at the 1% level

hub participation, suggesting that having additional sources of income reduces a household’s ability or interest in participating in dairy hubs. Milk production can be labour intensive, and in the case of resource constraints, farmers can hardly hire out labour for milk production. Thus, labour is important. Household size negatively affects dairy hub participation. This implies that as the number of household size increases, either less land becomes available for milk production or the family tends to be resource constrained for participation in dairy hub.

As indicated in the empirical specification, the identification of the model requires that at least one variable in the selection equation would not appear in the outcome equations. Distance to agriculture extension office is used as an identifying instrument. It is expected that distance to agriculture extension office will influence participation decision but will not directly affect milk yield or net returns. The distance to agriculture extension office residual estimates are not statistically significant, suggesting that the coefficients of the distance to agriculture extension office variable have been consistently estimated (Wooldridge 2015).

The covariance terms  $\rho_P$  and  $\rho_{NP}$  of the ESR model for the participants, as shown in Tables 2 and 3, are all statistically significant, indicating that the selection bias due to

unobservable factors occurred in participation. Thus, dairy hub participation may not have the same effect on the non-participants if they choose to participate. Hence, the use of ESR, which accounts for both observable and unobservable factors, is appropriate in this study (Lokshin and Sajaia 2004).

The statistically insignificant estimate for non-participants suggests that in the absence of dairy hub, there would be no significant difference in the average behaviour of the participants and non-participants caused by unobservable factors (Abdulai and Huffman 2014; Kleemann et al. 2014; Lokshin and Sajaia 2004). The negative signs for  $\rho_D$  indicate positive selection bias, suggesting that farmers with above-average yields and net returns have a higher probability of participating in dairy hub. Thus, comparative advantage tends to play a role in the determination of yields, net returns, and participation decision. The likelihood ratio tests for joint independence of the equation show that the equations are dependent.

### **Impact of dairy hubs on milk yields and net returns**

The estimates in the outcome equations in the columns for participants and non-participants in Tables 2 and 3 show the impacts of dairy hubs on milk yields and net returns. The results indicate that education is an important factor in explaining higher milk yields and higher net returns among the dairy hub participants. Educated farmers tend to have greater ability to decode new information and analyse the importance of innovations (Kassie et al. 2011). The positive and significant coefficient of the variable suggests that good knowledge and firm understanding of dairy hubs may increase the benefits of participating in dairy hubs in terms of milk yield and net returns.

Contact with extension agents has positive and significant impacts on milk yields and net returns for both participants and non-participants. Agricultural extension is the system of learning and building the human capital of farmers by giving information and exposing them to technologies that can increase agricultural productivity and welfare (Asfaw et al. 2012). Farmers who are frequently visited by government and non-government extension agents tend to be more progressive.

Access to credit tends to have a positive effect on milk productivity and net returns for both participating and non-participating farmers. In most cases, participation in dairy hub requires more investment cost; hence, the availability of credit to smallholder farmers who often have insufficient capital will encourage higher productivity and net returns. Landholding appears to have different impacts on adopters and non-adopters. The negative and significant coefficients for non-participants indicate that for this group of farmers, larger farmers obtained significantly lower yields and net returns than smaller farms. This finding supports the inverse farm size-productivity relationship, which posits that small farms are more productive than large farms (Chen et al. 2011). For participants, landholding did not significantly influence yields or net returns.

TLUs, the major household wealth proxy variable, have positive and significant impacts on yields and net returns of participants but no significant impact for non-participants. More livestock increases the availability of farmer's capital to dairy management. As argued by Langyintuo and Mungoma (2008), this result seems to suggest that relaxing an apparent liquidity binding constraint among poorly endowed households through access to wealth accumulation that shifts the household's wealth status to a

higher level will significantly increase their milk yields or net returns. Distance to dairy hub, which reflects remoteness, travel time or distance, is significant and negatively associated with milk yields and net returns of adopters and non-adopters. This might reflect both poorer information flow to more remote locations and weaker profit incentives for dairy hub among farmers who are less integrated into the dairy hub trading system. Due to higher input and lower output prices, traders in dairy hub channels tend to be sensitive to distance-related transaction costs and seek producers near paved roads or with their own transport.

Participation in non-farm activities tends to have a significant and positive effect on productivity and net returns of non-participants but no impact for participants. To the extent that non-participants more constrained compared to participants, the significant impact of non-farm activities on non-participants may be due the fact that off-farm earnings are used to purchase productivity-enhancing inputs like capital, labour and feeds to increase yields. Household size as proxied by the number of worker family members has a positive effect on milk yields and net returns, albeit statistically significant only for participants. This significant positive effect shows how family labour is important in developing countries, where moral hazard associated with hired labour is common (Asfaw et al. 2012).

**Estimating treatment effects (ATT)**

The estimates from the ATT, which show the impact of adoption on milk yields and net returns, are presented in Table 4. Unlike the mean differences presented in Table 1, which do not account for confounding factors, these ATT estimates account for the other confounding factors including selection bias arising from potential systematic differences between participants and non-participants. These results reveal that participation significantly increases yields and net returns. Specifically, the expected milk output per cow per day from dairy hub participation is 6.17 l compared to 4.67 from non-participants, representing a causal effect increase in output from participation by 32%. Similarly, participation increased net returns by 45%. These findings are consistent with the view that adoption of new agricultural technologies can improve farm productivity and household net returns (Abdulai and Huffman 2014; Di Falco et al. 2011)

**Conclusions**

This paper analyses the determinants and welfare impacts of participating in dairy hub in Kenya using recent data from a cross-section of farm households. Comparisons of average yields and net returns between participants and non-participants of dairy hubs revealed some significant differences. However, knowledge of average differences is not sufficient to understand participation decision across a sample of farmers, since they do

**Table 4** Impact of dairy hub participation on milk yields and net returns

	Mean outcome		ATT	t value
	Participants	Non-participants		
Milk yields	6.17	4.67	1.5***	6.27
Net returns	571	395	176***	5.41

ATT average treatment effect on the treated

\*\*\*Coefficient significant at 1% level

not account for the effects of other characteristics including factors influencing participation. Therefore, participation is modelled as a selection process, where the expected benefits to the dairy hub drive farmers' decisions. Given that farmers self-select themselves into participants and non-participants, an ESR approach is employed to account for selectivity bias and capture the differential impact of participation in dairy hub participants and non-participants.

The results showed that sample selection bias would have occurred if the outcome equations were estimated without accounting for observable and unobservable factors in the participation decision process. Thus, dairy hub participation may not have the same effect on the non-participants if they participated. Furthermore, a positive selection bias was observed for both milk yields and net returns, suggesting more productive farmers tend to participate in dairy hubs. Hence, comparative advantage tends to play a critical role in the determination of yields, net returns and participation decision.

On the impact of dairy hub participation on farm outcomes, the results showed a positive and significant relationship between participation and milk yields and farm net returns. In particular, participating in dairy hubs tends to increase milk yields by 32% and net returns by 45%, confirming the widely held view that productivity-enhancing agricultural innovations can contribute to increasing incomes of farm households and yields in developing countries (Becerril and Abdulai 2010). Developing mechanisms to help extend the dairy hubs to areas with high poverty rates is, therefore, a reasonable policy instrument to raise incomes in these areas, although complementary measures are needed. Policies encouraging the participation in dairy hubs through (for example, better support by extension agents or the provision of credit at affordable interest rates) should be supported by strategies to improve access to credit and encourage savings in rural areas in order to facilitate asset accumulation.

The findings also indicate a positive and significant influence of off-farm work on participation as well as impact on yields and net returns, confirming the importance of enhancing the ability of farm households to participate in non-farm work as a reasonable strategy to improve productivity and net returns in rural areas. Promising policy measures that can help boost non-farm work availability include increasing the access of rural households to assets such as financial capital and non-price factors such as education and infrastructure (Owusu et al. 2011). However, it is significant to mention that development of non-farm activities should complement the effort to develop agriculture, since activities in the former depend directly or indirectly on the later.

The positive influence of access to credit suggests that policies that enhance farmers' access to credit would go a long way to facilitate the participation in dairy hubs and increase milk yields and farm income. Hence, policy makers need to help farmers overcome financial and information barriers that are crucial in enhancing their participation in dairy hubs.

**Authors' contributions**

The author read and approved the final manuscript.

**Competing interests**

The author declares that he has no competing interests.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 25 July 2017 Accepted: 25 April 2018

Published online: 15 May 2018

## References

- Abdulai A, Huffman W (2014) The adoption and impact of soil and water conservation technology: an endogenous switching regression application. *Land Econ* 90(1):26–43
- Asfaw S, Shiferaw B, Simtowe F, Lipper L (2012) Impact of modern agricultural technologies on smallholder welfare: evidence from Tanzania and Ethiopia. *Food Policy* 37(3):283–295
- Becerril J, Abdulai A (2010) The impact of improved maize varieties on poverty in Mexico: a propensity score-matching approach. *World Dev* 38(7):1024–1035
- Chen Z, Huffman WE, Rozelle S (2011) Inverse relationship between productivity and farm size: the case of China. *Contemp Econ Policy* 29(4):580–592
- Deb P, Trivedi PK (2006) Specification and simulated likelihood estimation of a non-normal treatment-outcome model with selection: application to health care utilization. *Econ J* 9(2):307–331
- Di Falco S, Veronesi M, Yesuf M (2011) Does adaptation to climate change provide food security? A micro-perspective from Ethiopia. *Am J Agric Econ* 93(3):825–842
- Fuglie KO, Bosch DJ (1995) Economic and environmental implications of soil nitrogen testing: a switching-regression analysis. *Am J Agric Econ* 77(4):891–900
- Kassie M, Shiferaw B, Muricho G (2011) Agricultural technology, crop income, and poverty alleviation in Uganda. *World Dev* 39(10):1784–1795
- Kassie M, Zikhali P, Manjur K, & Edwards, S. (2009). Adoption of sustainable agriculture practices: evidence from a semi-arid region of Ethiopia. Paper presented at the Natural Resources Forum
- Khonje M, Manda J, Alene AD, Kassie M (2015a) Analysis of adoption and impacts of improved maize varieties in eastern Zambia. *World Dev* 66:695–706
- Khonje, M., Mkandawire, P., Manda, J., & Alene, A. (2015b). Analysis of adoption and impacts of improved cassava varieties. Paper presented at the 2015 Conference, August 9–14, 2015, Milan
- Kilelu, C. W., Klerkx, L., & Leeuwis, C. (2016). Supporting smallholder commercialisation by enhancing integrated coordination in agrifood value chains: experiences with dairy hubs in Kenya. *Experimental Agriculture*, 1–19
- Kleemann L, Abdulai A, Buss M (2014) Certification and access to export markets: adoption and return on investment of organic-certified pineapple farming in Ghana. *World Dev* 64:79–92
- Langyintuo AS, Mungoma C (2008) The effect of household wealth on the adoption of improved maize varieties in Zambia. *Food Policy* 33(6):550–559
- Lokshin M, Sajaia Z (2004) Maximum likelihood estimation of endogenous switching regression models. *Stata J* 4:282–289
- Maddala GS (1986) Limited-dependent and qualitative variables in econometrics. Cambridge University Press, Cambridge
- Makoni N, Mwai R, Redda R, van der Zijpp A, van der Lee J (2013) White gold: opportunities for dairy sector development collaboration in East Africa (pp. 1–203). Centre for Development Innovation, Wageningen UR (University and Research centre)
- Mojo D, Fischer C, Degefa T (2017) The determinants and economic impacts of membership in coffee farmer cooperatives: recent evidence from rural Ethiopia. *J Rural Stud* 50:84–94
- Muchichu RN (2014) The sustainability of dairy development in Tanzania: adoption of a participatory market chain approach system, MSc thesis in Agribusiness. University of Queensland, Queensland
- Mutinda G, Baltenweck I, Omondi I (2015) Setting up sustainable dairy business hubs: a resource book for facilitators. International Livestock Research Institute (ILRI) (aka ILCA and ILRAD), Nairobi
- Omondi, I., Zander, K., Bauer, S., & Baltenweck, I. (2016). Understanding farmers' preferences for artificial insemination services provided through dairy hubs. *animal*, 1–10
- Owusu V, Abdulai A, Abdul-Rahman S (2011) Non-farm work and food security among farm households in Northern Ghana. *Food Policy* 36(2):108–118
- Pyburn R, Eerdewijk A v (2016) A different kettle of fish? Gender integration in livestock and fish research. LM Publishers, Volendam
- Rademaker CJ, Bebe BO, van der Lee J, Kilelu C, Tonui C (2016) Sustainable growth of the Kenyan dairy sector: a quick scan of robustness, reliability and resilience. Wageningen University & Research, Wageningen Livestock Research, Wageningen
- Rao, E., Mtimet, N., Twine, E., Baltenweck, I., & Omoro, A. (2015). Cattle keepers' preference for dairy business hub options in Tanzania
- Rao E, Mtimet N, Twine E, Baltenweck I, Omoro A (2016a) Farmers' preference for bundled input-output markets and implications for adapted dairy hubs in Tanzania—a choice experiment. Paper presented at the 2016 AAAE Fifth International Conference, September 23–26, 2016, Addis Ababa
- Rao EJ, Omondi I, Karimov AA, Baltenweck I (2016b) Dairy farm households, processor linkages and household income: the case of dairy hub linkages in East Africa. *International Food and Agribusiness Management Review* 19(4):95–108
- Rao EJ, Qaim M (2011) Supermarkets, farm household income, and poverty: insights from Kenya. *World Dev* 39(5):784–796
- Tambo JA, Wünscher T (2017) Farmer-led innovations and rural household welfare: evidence from Ghana. *J Rural Stud* 55:263–274
- Vernooij, D. (2016). The Kenyan dairy value chain: promoting inclusive and climate smart dairy production
- Wooldridge JM (2015) Introductory econometrics: a modern approach: Nelson education
- Wossen T, Abdoulaye T, Alene A, Haile MG, Feleke S, Olanrewaju A, Manyong V (2017) Impacts of extension access and cooperative membership on technology adoption and household welfare. *J Rural Stud* 54:223–233