International Food and Agribusiness Management Review
Volume 19 Issue 4, 2016; DOI: 10.22434/IFAMR2014.0177

Received: 2 December 2014 / Accepted: 3 October 2016

Dairy farm households, processor linkages and household income: the case of dairy hub linkages in East Africa

RESEARCH ARTICLE

Elizaphan J.O. Rao[a], Immaculate Omondi[b], Aziz A. Karimov[c], and Isabelle Baltenweck[d]

[a]Agricultural economist, [b]MLE scientist, and [d]Project leader, International Livestock research Institute (ILRI), P.O. Box 30700, 00100, Nairobi, Kenya

[c]Agricultural economist, International Maize and Wheat Improvement Center (CIMMYT), Sehit Cem Ersever Caddesi 9-11, 06511 Ankara, Turkey

Abstract

In this study we have analysed the effects of household linkages to milk market via dairy hubs currently implemented under the East African Dairy Development project. Our analyses show that participation in dairy hubs increases dairy revenues by USD 1,022 on average. Impacts are higher for households participating in hubs supplying exclusively to processors (USD 1,673) relative to ones supplying hubs that pursue mixed-linkage approach. Moreover, participation in dairy hubs also yields significant effect on household income. Appropriate measures should be undertaken to widen the reach of such processor linkages while also safeguarding existing gains, more so as the processing sector becomes more concentrated.

Keywords: smallholder dairy farmers, large processors, dairy hubs, East Africa, dairy revenues, propensity score matching
JEL code: D4, L11, O13, Q12, Q13

[1]Corresponding author: J.Rao@cgiar.org

1. Introduction

Dairy production remains an important livelihood option for many poor rural households in the developing world; providing an important source of nutrients and contributing to household income (Duncan *et al.*, 2013; Thorpe *et al.*, 2000). In most developing countries, however, dairy value chains are characterized by multiple market failures that impede participation by dairy producers (Amorim *et al.*, 2013; Barrett, 2008; Wiggins *et al.*, 2010). First, majority of dairy producers in the developing world are smallholders producing low volumes (Holloway *et al.*, 2000; Thornton and Herrero, 2001). This coupled with the scattered nature of their location makes them unattractive suppliers to more structured and reliable market outlets such as processors. Moreover, these poor households are also remotely located with limited access to reliable infrastructure, which leads to higher transaction costs, further compromising their ability to access structured markets (Jayne *et al.*, 2010). Limited access to input markets also heighten cost of production further restricting households to low-input-low-output vicious cycle.

In response to these limitations, development agencies continue to promote approaches aimed at enhancing market participation by smallholder dairy producers. These initiatives include provision of market information and promotion of collective action as a means of enhancing access to both input and output markets (Fischer and Qaim, 2012; Njuki *et al.*, 2011). It is expected that with better linkages to markets producers would realize higher and sometime less volatile output prices and lower transaction costs in accessing inputs leading to higher net returns from dairy production. This would lead to improved income for participating households and possibly enhance access to inputs and improved technologies.

One such initiative is the dairy hub model implemented under the East Africa Dairy Development (EADD) project that was initiated in 2008 and is currently in its second phase (2013 to 2018). Dairy hubs are geared towards upgrading dairy value chains via linkages to input and output markets mainly through collective action. Through aggregation and bulk selling of milk farmers accrue bargaining advantage when negotiating with milk buyers. This is likely to have positive effect on milk prices for participating farmers. Similarly, households benefit from bulk sourcing of inputs and collectively negotiated rates with service providers thus lowering the cost of production and subsequently leading to improved dairy and household income. The EADD project also provides capacity support to hubs with an aim of making the supported hubs self-sustaining by the end of the project's implementation period – the year 2018.

Regarding milk sales, dairy hubs pursue diverse market strategies including: exclusive sales of milk to processors and a mixed strategy involving sales to processors and local consumer outlets. In an attempt to understand the impacts of this development initiative, an evaluation was carried out at the end of phase one of EADD and findings showed improved welfare for participating households in terms of household income and dietary diversity score. This was true across Kenya, Uganda and Rwanda where the project was implemented. However, the study largely adopted descriptive approaches with less statistical rigor. Moreover, the analyses did not expressly look at impacts of different marketing strategies on returns to dairy production.

In this study we analyse the impacts of different marketing strategies adopted by dairy business hubs on dairy and household income. In particular, we evaluate whether linkages to processors have greater impacts on household welfare. Using data collected from smallholder dairy farm households living within the catchment areas of dairy hubs in Kenya and Uganda (dairy hubs participants and non-participants) and a mix of descriptive analyses and propensity score matching approaches, we provide evidence on the market linkage mechanism that yields the greatest impact on dairy and household income. In the next section we present the analytical framework guiding this study. We then present data and descriptive statistics. This is followed by a presentation and discussion of the results before providing concluding remarks on possible policy implications.

2. Analytical framework

Program evaluation often follows approaches suggested by Maddala (1983):

$$y = X\beta + \gamma I + u \tag{1}$$

Where y can be considered as household or dairy income or any other household welfare indicator; *X* is a vector of farm, household and contextual characteristics that could influence dairy/household income; and I is a dummy indicating whether or not a household participates in a hub. Holding other factors constant then, the coefficient (γ) captures partial effects of household participation in a hub (in general; that are processor-linked or; hubs linked to other market outlets) on dairy revenue/household income. However, because dairy producers may self-select into participation in hubs, this estimate may be biased. In other words, it is possible that some determinants of hub participation may also affect dairy revenue/household income. If such factors are not included explicitly in Equation 1, as is the case when such variables are unobserved, then the indicator for hub participation in Equation 1 will be correlated with the error term (u) leading to a biased estimation of γ. If participation were randomized, the counterfactual would be observable, making it possible to derive causal inference. Unfortunately, this is not the case in our example. The cross-sectional nature of our data also rules out the possibility of addressing selection bias through panel data approaches.

To address the potential selection bias, we propose a matching technique, which assume that conditioning on observable variables eliminates sample selection bias (Heckman and Navarro-Lozano, 2004). Similar approach has been used in the context of agricultural technology adoption (Ali and Abdulai, 2010; Faltermeier and Abdulai, 2009). Matching models essentially create an experimental condition in which participation in dairy hubs is randomly assigned, thus allowing for identification of causal link between hub participation and dairy/household incomes.

We use a class of matching models known as propensity score matching (PSM) to measure the effects of participation in processor-linked hubs of various types on household dairy revenue/household income. Instead of directly comparing dairy revenues or household income between households participating in dairy hubs (processor-linked or otherwise) and their counterparts not participating in these hubs, PSM compares between only households participating in dairy hubs ('treated') and those households not participating in dairy hubs ('control'). Moreover, PSM only compares 'treated' and 'control' households that are similar in terms of observable characteristics, thus reducing the bias that would otherwise occur if the two groups are systematically different (Dehejia and Wahba, 2002). PSM involves two stages. In the first stage, we generate propensity scores P(z) from a probit model that estimates the probability that a household participates in a market linkage program (processor-linked hub for instance). The vector z is of observed conditioning variables that may overlap with variables included in *X* in Equation 1. We then construct a control group by matching participants in the hub arrangement with participants in other market linkage programs based on similarity of their propensity scores. Households in 'control' group for whom appropriate matches cannot be found as well as those not used as matches are dropped. In the second stage, we calculate the average treatment effect on the treated (ATT) households for the outcome variable (household and dairy income), using matched observations of households in the 'treated' and 'control' groups. The PSM estimator of the ATT is the difference in outcomes between treatment and control groups, appropriately matched by the propensity scores:

$$\tau_{ATT}^{PSM} = E_{\mathrm{P(z}\,|\,I=1)}\,[E\{R_1 \,|\, I=1, P(z)\} - E\{\mathrm{R}_0 \,|\, I=0, P(z)\}] \tag{2}$$

where R_1 and R_0 are outcomes for the treated and control farms respectively; $I = 1$ indicates treated households and $I = 0$ control households.

There are various matching techniques, but the most common ones include nearest neighbour matching (NNM), kernel-based matching (KBM), stratified radius matching, and Mahalanobis matching (Caliendo and Kopeinig, 2008). In this study, we apply the KBM and the NNM methods. NNM involves pairing farmers in processor-linked hubs and non-processor linked hubs that are closest in terms of P(z) as matching partners. KBM, on the other hand, uses a weighted average of the outcome variable for all individuals in the control group (households in non-processor linked hubs) to construct a counterfactual outcome. Observations that provide better matches are given more weight. The weighted average is compared to the outcome for households in processor-linked hubs, and the difference provides an estimate of the treatment effect for each household supplying the processor-linked hub. A sample average over all processor linked households then provides an estimate of ATT.

It is worthwhile noting that PSM can control only for selection bias that is due to observed factors z. In other words, systematic differences between processor linked farmers and non-processor linked farmers may still exist even after conditioning, especially if part of the selection process is based on unobserved variables (Smith and Todd, 2005). Our estimation of ATT is based on the assumption that the distribution of such unobservables is the same for treatment and control groups. However, this is ultimately an empirical question that should be tested (Imbens, 2004). Therefore, we apply the standard bounding test proposed by Rosenbaum (2002), which evaluates how strongly the unobserved variables would have to influence selection to invalidate the implications of the matching process.

3. Data and descriptive statistics

Data

Dairy farming households in EADD-supported hubs catchment areas were surveyed as part of a baseline for the second phase of EADD project (EADD phase II). Two performance indicators for EADD, increase in milk production and income from milk production, were the main response variables used in estimating the required sample size. Due to the large number of hubs in Uganda, stratification by cattle production system (intensive versus intensive/semi-intensive) was done, for Uganda, in order to estimate the required sample size that is sufficient to elicit the desired response. Consequently, all the dairy hubs supported by phase II of EADD in Kenya (8) and a sample of 24 hubs (out of 33) in Uganda, were included in a household survey conducted between October and December 2014. The required sample sizes for the two countries were estimated to be 322 and 671 cattle keeping farm households from 8 EADD-hubs in Kenya and 24 EADD-hubs Uganda respectively, with equal sample sizes per hub in each country.

Geo-spatial random sampling technique was used to randomly select smallholder dairy farm households living within catchment area of each dairy hub in Kenya and Uganda. Using a structured questionnaire, data was collected from these households through personal interviews. In Kenya, the 8 hubs' catchment area covered Nandi East, Nandi North, Nandi south, Sotik, Narok South, Trans Nzoia and Wareng districts in North and South Rift (Western Kenya). In Uganda, the hub catchment areas span across 13 districts, i.e. Isingiro, Ibanda, Kiruhura, Ssembabule, Masaka, Mukono, Jinja, Kayunga, Kamuli, Kyankwanzi, Wakiso, Kiboga and Mityana districts. Socio-demographic data, data on livestock assets, milk production and utilization, and input use were collected from sampled households in the two countries. In addition to the household survey data, information regarding market outlets used by the dairy hubs, their linkages with processors and other market outlets, were collected between September 2014 and January 2015, from monthly business reports submitted to EADD by the hubs. A sustainability assessment study, conducted between March and April 2015 was used to obtain information on levels of hub sustainability.

Dairy hubs and marketing strategies

Dairy hubs are farmer-owned collective action-based mechanisms for enhancing market linkages for smallholder dairy farmers. In a situation where smallholder producers are scattered and produce low volumes, it is uneconomical for milk traders and/or processors, as well as input and business service providers to deliver services to farmers. Through bulking and/or chilling, the hubs enable farmers to supply milk to large dairy processors who are the main players in the dairy output market. Courtesy of collective action dairy hubs also command large demand for inputs and services, which could be attractive to inputs and service providers. Besides, most hubs implement a check-off system that enables farmers to access inputs and services on the account of milk delivery, which allows households to access inputs and services even when they do not have cash. Among other services, hubs also provide farmers easy access to loans and training on animal husbandry, which in turn improves smallholder farm management. Hubs therefore reduce transaction costs both for suppliers and buyers of milk thus improving margins from dairy production. Processors on the other hand benefit from reliable supply of high-quality local milk and achieve better control over the supply chain. Finally, local/regional consumers realize gains from affordable and safe milk delivered via efficient milk marketing mechanism and better regulated milk supply.

While development partners have been supporting this initiative, hubs remain purely farmer-owned with development partners only playing a facilitative role. This is achieved through capacity building at both farmer and hub management level. Development partners also conduct studies to evaluate performance of the hubs both at farm and hub levels. Development partners also facilitate sharing of lesson that can enhance farm and hub level performance and thus move hubs towards sustainability.

With regards to milk marketing strategies, some hubs sell milk exclusively to processors (pure processor hubs) what we refer to here as ‘strategy 1’. On the other hand some hubs sell milk to diverse outlets with large processors being just one of the clients (mixed-linkage hubs) – ‘strategy 2’. Selling to more than one outlet is a risk managing strategy through which hubs may take advantage of potentially higher prices in non-processor outlets. However, non-processor prices are subject to wide fluctuations and this may erode gains from period of higher milk prices. Therefore depending on the share of hub’s milk going to non-processor outlet, farmers attached to such hubs may experience lower prices on average and hence lower annual revenues. On the other hand, while processor prices may be low, the option offer stable prices and hence more stable revenue flows for associated farmers throughout the year. Differences in market outlets therefore imply differences in profits that a hub can generate. We hypothesize that these differences would trickle down to farmers in terms of prices that farmer receive for milk sales, thus impacting on dairy and possibly household income. In order to understand the effect of linkage to large processors, this study compares smallholder dairy farmers participating in these different types of dairy hubs:

- First we compare households participating in dairy hubs versus those not participating in the hubs (Figure 1). By participation we mean those households that either deliver milk or access inputs and/or services via hub arrangement. Both elements of participation are expected to impact dairy income – milk sales via revenues and input/service access via cost of production. Non participants in this

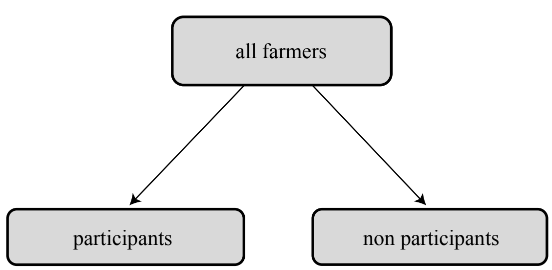

Figure 1. The first step compared participating households to those not participating in the hubs.

case are livestock keepers who are residents in the catchment areas of respective hubs but neither sell milk nor access inputs and/or services from the hub.

- We then compare participants and non-participants who reside within the catchment of dairy hubs supplying milk exclusively to large processors, i.e. hubs pursuing 'strategy 1' (Figure 2).
- In order to understand the differential impacts of different market linkages, we also compare participants and non-participants in the catchment areas of hubs pursuing 'strategy 2' (Figure 2).

Descriptive statistics

Table 1 shows a profile of dairy hubs categorized by country. Approximately 31,700 smallholder dairy farmers were registered as suppliers in the dairy hubs in the two countries by April 2015. These dairy hubs supplied processors and other market outlets with a total of 82,700 liters of milk per day. The dairy hubs in Kenya were all chilling plant hubs with 2 of them adopting a 'pure processor linkage' approach while 6 hubs had a 'mixed linkage' approach. On the other hand, 10 out of 24 dairy hubs in Uganda adopted a 'pure processor linkage' approach.

In Table 2, we show some selected socio-economic characteristics of smallholder dairy farmers categorized by participation in dairy farmer groups. We see that farmers actively participating in dairy hubs are significantly more educated. They also have significantly more experience with dairy farming. We also see that majority of farmers in our sample that participate in dairy hubs are from Uganda (58% of participants are from Uganda compared to Kenya's 42%). Finally we note that significantly more active participants are found in EADD-supported dairy hubs. Similarly significantly more active participants are in hubs linked to processors compared to non-active participants.

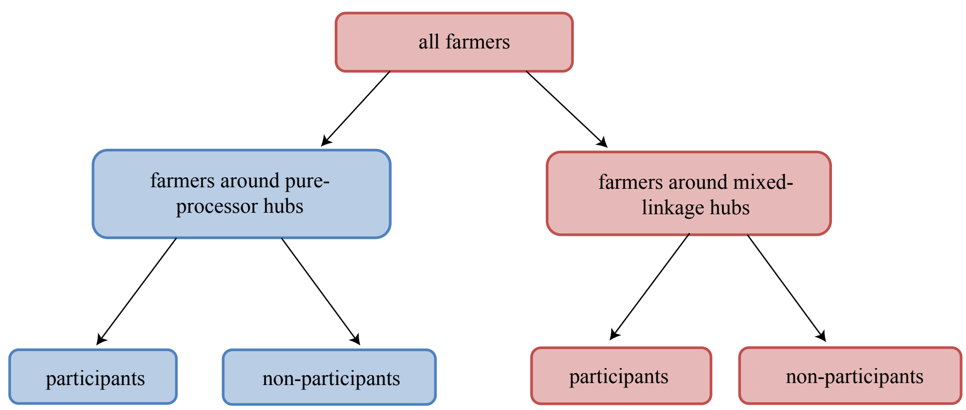

Figure 2. Participants and non-participants using mixed linkage hubs are compared.

Table 1. Description of the dairy hubs by country, marketing strategy and hub types.

Form of linkage	**Pure processor linkage**		**Mixed linkage**	
	Kenya	**Uganda**	**Kenya**	**Uganda**
number of dairy hubs	2	10	6	14
number of registered suppliers	6,507	1,483	21,534	2,230
volume of milk (liters) supplied to market outlets per day	11,435	41,007	21,805	8,539

Table 2. Selected socio-economic characteristics of small holder dairy farmers by country and participation in dairy farmer groups.[1]

	Participants n=193		**Non-active participants n=800**	
	Characteristic	**St. error**	**Characteristic**	**St. error**
age of operator	52	15	50	14
annual dairy revenues (USD)	1,640*	2,446	365	983
total household income (USD)	12,954*	26,288	5,379	14,731
education (years of schooling)	8*	5	7	5
dairy farming experience (years)	20*	14	16	12
proportion of male operators	83.4	37.3	83.1	37.5
farming as a primary occupation (%)	77.5	41.8	77.3	41.9
sample households from Kenya) (%)	42*	49	30	46
sample households in EADD supported groups (%)	70*	46	4	20
sample households in non-EADD supported groups (%)	4	19	5	21
sample households in processor-linked hubs (%)	36*	43	23	42

[1] Asterisks indicate significance at 1%.

Income distribution

We also present a breakdown of sources of household income in addition to average income comparison. The illustration in Figure 3 shows that households in the EADD regions engage in other income generating activities besides dairy. As expected, majority of households engage in sale of milk and other dairy products that are own produced. This is followed by the proportion of households generating income from sale of cattle and other livestock as well as livestock products. Crop revenues and trade in agricultural products (not own produced) follow next in that order.

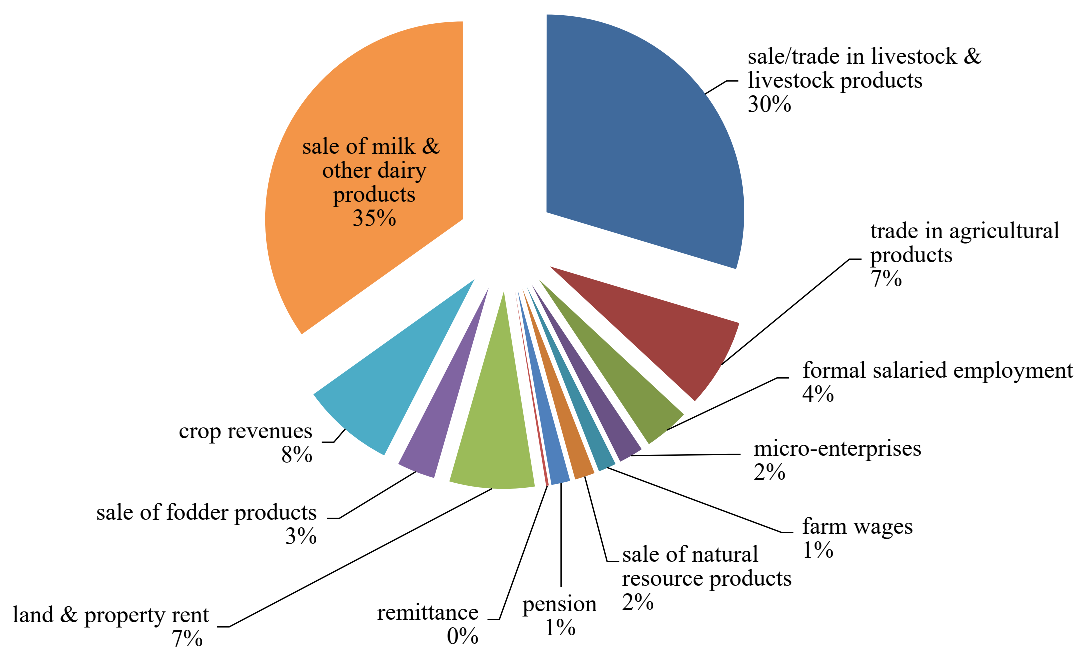

Figure 3. Proportion of households engaging in respective income generating activities.

We further illustrate in Figure 4 the average amount of revenue generated from respective sources. First, we see that households in the catchment of 'pure processor' hubs have higher household income than households in the catchment of 'mixed-linkage' hubs. For the whole sample, Figure 4 also shows that household income is largely generated from sale and trade in livestock and livestock products. This is also true for households residing in the catchment of dairy hubs pursuing 'mixed linkage' approach. As for households residing in the neighbourhood of 'pure processor' dairy hubs, micro-businesses provide the highest contribution to household income. Finally, while dairy revenue ranks low relative to income from other sources, households residing in the catchment of 'pure-processor' dairy hubs tend to generate higher revenues from this source than their counterparts in the 'mixed-linkage' hubs.

4. Empirical results and discussions

Descriptive results discussed in the previous section reveal some differences in dairy and household income between active participants and non-active/non-participants in dairy hubs. However, it is impossible to determine if these differences are due to household participation in respective market linkage programs. In order to attribute these differences to hub participation we conduct statistical matching as described in the analytical framework. Results of these analyses are shown in Table 3. In Supplementary Table S1 we also show results of the probit model used to predict propensity scores that form the basis for our matching.

The statistical matching results shown in Table 3 reveal that participation in EADD dairy hubs leads to significantly higher revenues from dairy production. For the whole sample, we find that holding all factors constant, participation in dairy hubs increases annual dairy revenue by USD 1,022 on average. The analyses also show that participation in dairy hubs has a positive and significant effect on total household income – increasing total annual household income by USD 4,628 on average. The larger effect on total household income implies a multiplier effect of dairy revenue; dairy revenues may actually be used as capital in other household income generating activities thus enlarging the total effect of these revenues on household income. For instance, dairy revenues could be supporting micro-businesses which as Figure 4 revealed, is a major contributor to household income especially for participants in the 'pure processor' hubs. These results are irrespective of the marketing strategy adopted by the dairy hubs.

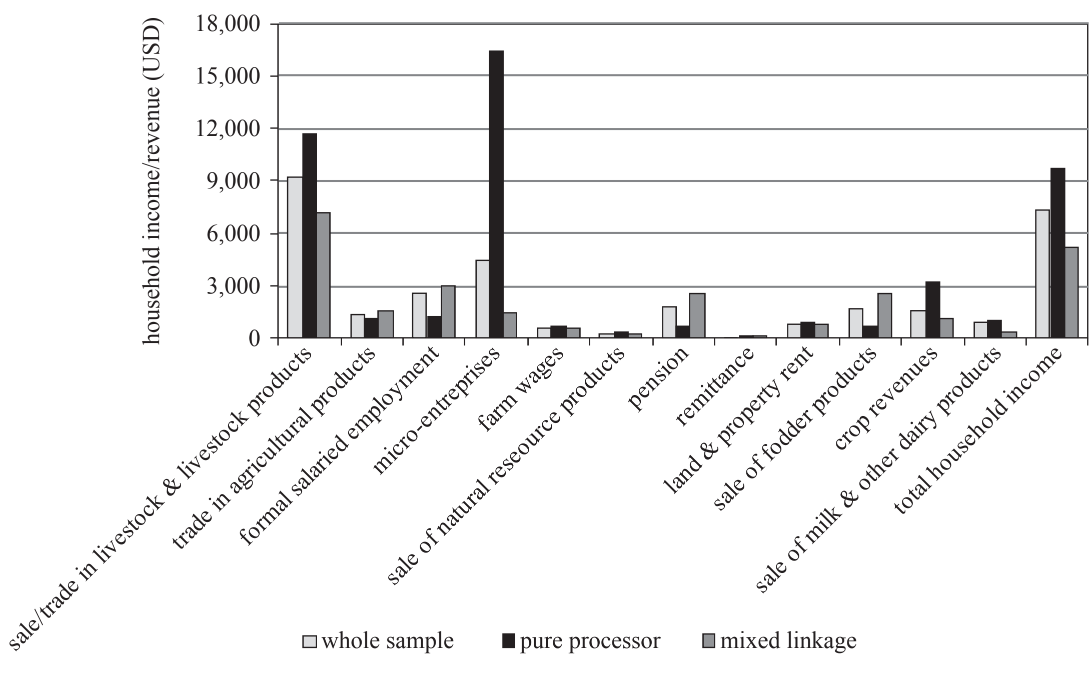

Figure 4. Average household earnings from different activities.

Table 3. Results of the propensity score matching.

Matching algorithm	Outcome	ATT[1]	z-statistic[2]	Critical level of hidden bias (Γ)	Treated (n)	Control (n)
Effect of household participation in dairy hubs						
nearest neighbour	annual dairy revenue	1,022.81	4.26***	2.00-2.05	186	738
	total household income	4,628.69	2.13**	1.35-1.40	186	738
kernel	annual dairy revenue	990.31	4.62***	1.60-1.65	186	738
	total household income	4,690.60	2.12**	1.20-1.25	186	738
Effect of household participation in pure processor-linked dairy hubs						
nearest neighbour	annual dairy revenue	1,673.22	2.64***	2.55-2.60	97	242
	total household income	7,463.28	1.34	–	97	242
kernel	annual dairy revenue	1,676.41	3.02***	2.10-2.15	97	242
	total household income	7,682.69	1.59	–	97	242
Effect of household participation dairy hubs with mixed-linkage						
nearest neighbour	annual dairy revenue	747.41	2.81***	3.60-3.65	89	494
	total household income	3,807.36	1.27	–	89	494
kernel	annual dairy revenue	654.10	2.72***	2.50-2.55	89	494
	total household income	3,216.00	1.38	–	89	494

[1] ATT = average treatment effect on the treated.

[2] ** and *** are significant at the 5 and 1% levels, respectively. The z-values for the ATTs are based on bootstrapped standard errors with 500 replications.

We also note that while hubs are supposed to enhance access to inputs and services and could thus lower production costs for households, we believe that higher dairy revenues for participating households are largely due to higher prices since as can be seen in Figure 5, dairy revenues are largely driven by milk sales.

When we narrow our focus to only those households in the catchment of hubs supplying exclusively to processors, the treatment effect of participation in dairy hubs on dairy revenue is even larger. For this category of households, participation in dairy hubs yields an impact on annual dairy revenues to the tune of USD 1,673. These findings are in contrast with Navarro *et al.* (2015) which found for the case of Peru that informal markets tend to offer higher profits per litre of milk than formal channels. Similarly, average treatment effects on treated for household income are larger for this sub-sample of households. This effect is,

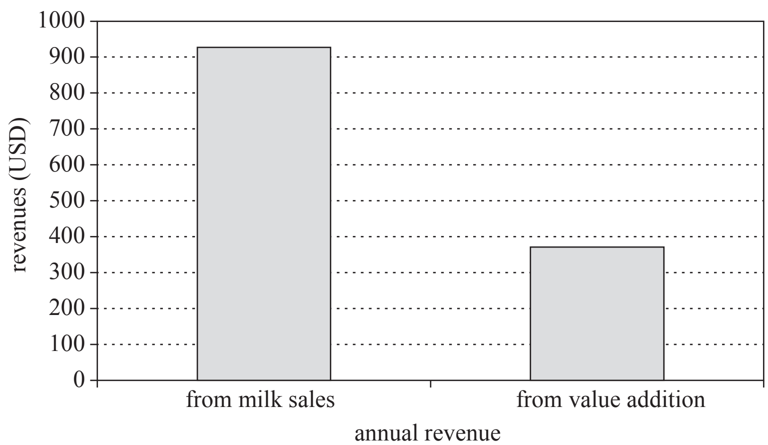

Figure 5. Components of dairy revenues.

however, insignificant. This is possibly due to the value of dairy revenue relative to total household income. As shown in Figure 4, dairy revenue for households in pure processor linkage is fairly low compared to total household income (1000=5 USD to 9,742 USD). This is comparison to the whole sample where the gap is relatively not wide (964 USD to 7,292 USD).

We make a further comparison between participants and non-participants for households in the catchment of 'mixed-linkage' hubs. The lower panel of Table 3 shows that participation in dairy hubs for this sub-sample of households also yields positive and significant effects on dairy revenues, albeit by far lower magnitude relative to 'pure processor' linkage. Participation in this category of hubs leads to an increase in dairy revenues by 747 USD on average. As already explained, while price offers by processors may be low relative to alternative outlets, stability in prices ensures that overall annual returns are higher for households supplying pure processor-linked hubs. This is in comparison to mixed linkage approach where wide fluctuations in price offers by non-processor outlets may erode gains from windows of higher prices leading to low prices on average and hence low annual revenues. Treatment effects on household income are, however, insignificant for this sub-sample. Finally, we test for statistical difference in the distributions of treatment effects between the two sub-samples. Figure 6 compares the distribution of the two treatment effects (for 'pure processor' and 'mixed linkage' sub-samples). Kolmogorov-Smirnov test confirms that the cumulative distribution function (CDF) of treatment effects for 'pure processor' linkage statistically dominates the CDF of treatment effects for 'mixed-linkage'.

Sensitivity analyses

The main weakness of PSM relies on the fact that program participation is only explained by observed (observable) covariates. The approach would therefore be effective as long as bias from unobserved covariates remains minimal. However, if program participation variables that are usually used to balance the treated and comparison sub-samples are incomplete, PSM results can be biased. It is therefore critical that factors driving participation in the program (especially dropped nonparticipant are carefully investigated and are included in the modelling of PSM to the extent possible. Including more observed variables ensures that matched treated and untreated observations are as similar as possible. This follows from the assumption that some of the unobserved factors may be related to many matching variables included in the PSM modelling, which allows for reduction of the potential bias that emanated from omission of unobserved variables. We then test for bias reduction (BR) by estimating a bias reduction index as follows:

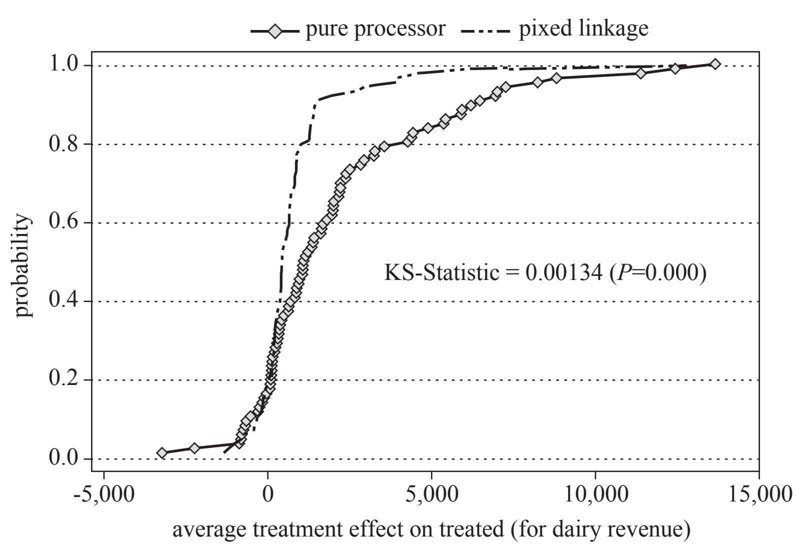

Figure 6. Cumulative distribution of treatment effects on dairy revenues by market linkage. KS = Kolmogorov-Smirnov test.

$$BR = 100\left(1-\frac{Bias_{after}}{Bias_{before}}\right) \tag{3}$$

BR values greater than 20% are considered large – indicating substantial reduction of bias achieved via matching (Rosenbaum and Rubin, 1985). Additionally pseudo R-squared should be fairly low after matching to ensure that there are minimal systematic differences in the distribution of covariates between treated and comparison groups, i.e. the two groups are more similar thus allowing for unbiased comparison of outcome variable.

Another issue with PSM is that it requires large sample of non-participants from which matches are drawn. It is needed so that enough variation is provided in the representative sample. Otherwise interpretation of treatment effect results will mislead policy implications.

We show results of this BR test in Table 4, which reveals that the variance of treatment status explained by covariates declined substantially after matching. Similarly likelihood ratio test (*P*-value) shows that the joint significance of covariates on treatment status cannot be rejected before matching, while it is rejected after matching. The joint insignificance of covariates together with the low pseudo R-squared after matching imply that there is no systematic differences in the distribution of covariates between groups after matching. Our matching is therefore based on fairly similar observations of matched and comparison groups with minimal bias if any.

While test results in Table 4 have shown that our matching procedure was successfully able to balance the distribution of observed characteristics, hidden bias may still arise if there are unobserved variables that simultaneously affect assignment into treatment. Matching estimators are not robust to such hidden bias. We therefore test for potential hidden bias from farmer heterogeneity due to unobserved variables using the bounding approach suggested by Rosenbaum (2002) and explained in the analytical framework. Assuming

Table 4. Indicators of covariate balancing before and after matching.[1]

Matching algorithm	Outcome	Median absolute bias (before matching)	Median absolute bias (after matching)	% bias reduction	Pseudo R (unmatched)	Pseudo R (matched)	*P*-value of LR (unmatched)	*P*-value of LR (matched)
Whole sample								
nearest neighbour	annual dairy revenue	32.9	8.5	74.2	0.55	0.05	0.000	0.503
	total household income	32.9	8.5	74.2	0.55	0.05	0.000	0.503
kernel	annual dairy revenue	32.9	8.2	75.1	0.55	0.04	0.000	0.770
	total household income	32.9	8.2	75.1	0.55	0.04	0.000	0.770
Processor-linked hubs								
nearest neighbour	annual dairy revenue	21.5	11.0	48.8	0.564	0.165	0.000	0.187
	total household income	21.5	11.0	48.8	0.564	0.165	0.000	0.187
kernel	annual dairy revenue	21.5	14.4	33.0	0.564	0.126	0.000	0.544
	total household income	21.5	14.4	33.0	0.564	0.126	0.000	0.544
Mixed-linkage hubs								
nearest neighbour	annual dairy revenue	47.2	21.8	53.8	0.682	0.241	0.000	0.163
	total household income	47.2	21.8	53.8	0.682	0.241	0.000	0.163
kernel	annual dairy revenue	47.2	12.2	74.2	0.682	0.108	0.000	0.960
	total household income	47.2	12.2	74.2	0.682	0.108	0.000	0.960

[1] LR = Likelihood ratio test.

two individuals have the same observed covariates z (as implied by the matching procedure), the two matched observations would differ in their odds of participating in the dairy hubs only by the difference in unobserved covariates, which is measured by the parameter Γ. The test procedure involves changing the level of Γ and deriving the bounds on the significance levels of the ATT under the assumption of endogenous self-selection into dairy hub participation. This allows for identification of the critical levels of Γ at which the estimated ATT would become insignificant.

Results of this test are shown in the fifth column of Table 3. Using the example of dairy revenues for the pure processor linkage, the critical values for hidden bias (Γ) are 2.10-2.15 with KBM and 2.55-2.60 with NNM. The lowest value of Γ=1.6 implies that individuals that have the same z-vector would have to differ in their odds of participation in dairy hubs by at least a factor of 1.6 (60%) in order to render the ATT for dairy revenues insignificant. Even though unobserved variables may play a certain role, it is very unlikely that they would influence the odds of participation in dairy hubs to such a great extent.

5. Conclusions

Dairy activities account for a significant proportion of household income in the milk producing zones. However, potential for increased dairy incomes is compromised by many smallholders operating below capacity, largely occasioned by limited access to services, essential inputs and business support. Consequently, milk prices are usually low while input costs are high, thus constraining milk profit margins. These limitations have motivated various market linkage mechanisms aimed at increasing participation by households in output markets while also linking households to input and service markets. Dairy business hubs is one such imitative currently implemented under the EADD project. By bulking and chilling milk through the hubs, farmers can bargain for higher prices from milk processing companies. Additionally, the hub offers a one-stop source of essential dairy-related inputs and services such as feeds, drugs, breeding, animal health and extension services, usually under flexible payment arrangements.

While milk prices may sometime be lower in hubs than other outlets, the possibility of accessing inputs and services cost effectively and under flexible payment arrangements makes hubs a preferred market access mechanism for many households. These gains translate to increased production levels and enable farmers to cost-effectively produce higher volumes of milk, thus leading to increased dairy income. The effects may also spill over to total household income. In this study, we have analysed the effects of dairy households' participation in the hub as implemented under the EADD project.

First, our findings show that participation by households in dairy hubs significantly increases both dairy revenues and household income, irrespective of the market linkage pursued by respective dairy hubs. Sub-sample analyses (market-linkage sub-samples) reveal even higher effects for households participating in dairy hubs that sell milk exclusively to processors. This is in comparison to households that supply milk to dairy hubs following a mixed marketing approach. Another important result relates to the multiplier effect of dairy income on total household income, possibly due to investment of returns from dairy in other revenue generating enterprises at household level.

Yet processors and other formal outlets often have difficulties buying milk from smallholders due to their scattered and remote locations. Appropriate measures need to put in place to eliminate some of the barriers that smallholder experience in trying to access processor outlets. Morgan (2009) underscores the need for collective approaches involving cooperatives as a means to accessing the more formal milk markets including processors. However, the success of such collective approaches hinges crucially on good governance, limited state management and the fit of respective collective action approaches to cultural and socio-economic context. More flexible approaches involving business-oriented farmer groups operating outside the tenets of cooperative laws could be more successful in linking smallholders to formal markets including processors. The EADD project recognizes this fact and has therefore encouraged formation of the flexible and business-oriented producer organizations. Producer organisations (POs) at the centre of dairy hubs decide on their market linkage strategy, whether to sell exclusively to processors or to diversify to different milk buyers, based on their internal business strategy as well as their external environment. Based on these results, development agencies and facilitators should be encouraged to work closely with POs to identify the most appropriate market linkage strategy, including focusing on dairy processors, to maximise impact on dairy farmers income.

In spite of the impressive effects of household participation in pure processor-based linkages, caution needs to be taken to avoid monopolistic tendency that emerge as consolidation occurs in the processing sector. Increased consolidation often leads to competition among large processors and occasionally to lower prices paid to producers (Morgan, 2009). To safeguard against such ills, farmer organizations/dairy hubs should pursue more formal contracts with processors stipulating purchase prices and payment schedules. Some more advanced dairy hubs could also integrate vertically by establishing their own processing units, which would give farmers even guarantee better prices.

Finally, in light of the apparent benefits and given the low levels of farmer participation dairy hubs currently, future work should seek to understand incentives/disincentives for farmer participation in the hub arrangements.

Supplementary material

Supplementary material can be found online at https://doi.org/10.22434/IFAMR2014.0177.

Table S1. Propensity score model.

Acknowledgements

This article is based on the research conducted on the East Africa Dairy Development program project and we thank all partners and donors. It was also partly funded by the Livestock and Fish program of the CGIAR and we therefore thank all donors that globally support our work through their contributions to the CGIAR system (http://www.cgiar.org/about-us/our-funders). We extend our appreciation to the project team and the dairy producers and Producers Organizations for their cooperation during the research. The usual disclaimers apply.

References

Ali, A. and A. Abdulai. 2010. The adoption of genetically modified cotton and poverty reduction in Pakistan. *Journal of Agricultural Economics* 61: 175-192.

Amorim, P., H. Meyr, C. Almeder and B. Almada-Lobo. 2013. Managing perishability in production-distribution planning: a discussion and review. *Flexible Services and Manufacturing Journal* 25: 389-413.

Barrett, C.B. 2008. Smallholder market participation: concepts and evidence from eastern and southern Africa. *Food Policy* 33: 299-317.

Caliendo, M. and S. Kopeinig. 2008. some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys* 22: 31-72.

Dehejia, R.H. and S. Wahba. 2002. Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics* 84: 151-161.

Duncan, A., N. Teufel, K. Mekonnen, V. Singh, A. Bitew and B. Gebremedhin. 2013. Dairy intensification in developing countries: effects of market quality on farm-level feeding and breeding practices. *Animal* 7: 2054-2062.

Faltermeier, L. and A. Abdulai. 2009. The impact of water conservation and intensification technologies: empirical evidence for rice farmers in Ghana. *Agricultural Economics* 40: 365-379.

Fischer, E. and M. Qaim. 2012. Linking smallholders to markets: determinants and impacts of farmer collective action in Kenya. *World Development* 40: 1255-1268.

Heckman, J. and S. Navarro-Lozano. 2004. Using matching, instrumental variables, and control functions to estimate economic choice models. *Review of Economics and Statistics* 86: 30-57.

Holloway, G., C. Nicholson, C. Delgado, S. Staal and S. Ehui. 2000. Agroindustrialization through institutional innovation transaction costs, cooperatives and milk-market development in the East-African highlands. *Agricultural Economics* 23: 279-288.

Imbens, G.W. 2004. Nonparametric estimation of average treatment effects under exogeneity: a review. *Review of Economics and Statistics* 86: 4-29.

Jayne, T.S., D. Mather and E. Mghenyi. 2010. Principal challenges confronting smallholder agriculture in Sub-Saharan Africa. *World Development* 38: 1384-1398.

Maddala, G.S. 1983. Limited-dependent and qualitative variables in econometrics. Cambridge University Press, Cambridge, UK.

Morgan, N. 2009. Models and opportunities for smallholder dairy producers in Asia: lessons learned. *Smallholder dairy development: lessons learned in Asia*. Available at: http://tinyurl.com/oxpjjnd.

Navarro, E.F., G. Faure, E. Cortijo, E. De Nys, J. Bogue, C. Gómez, W. Mercado, C. Gamboa and P.-Y. Le Gal. 2015. The impacts of differentiated markets on the relationship between dairy processors and smallholder farmers in the Peruvian Andes. *Agricultural Systems* 132: 145-156.

Njuki, J., S. Kaaria, A. Chamunorwa and W. Chiuri. 2011. Linking smallholder farmers to markets, gender and intra-household dynamics: does the choice of commodity matter&quest. *European Journal of Development Research* 23: 426-443.

Rosenbaum, P.R. 2002. *Observational studies*. Springer Science+Business Media, New York, USA.

Rosenbaum, P.R. and D.B. Rubin. 1985. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician* 39: 33-38.

Smith, J.A. and P.E. Todd. 2005. Does matching overcome Lalonde's critique of nonexperimental estimators? *Journal of Econometrics* 125: 305-353.

Thornton, P. and M. Herrero. 2001. Integrated crop-livestock simulation models for scenario analysis and impact assessment. *Agricultural Systems* 70: 581-602.

Thorpe, W., H. Muriuki, A. Omore, M. Owango and S. Staal. 2000. Dairy development in Kenya: the past, the present and the future. Paper presented at the Annual Symposium of the Animal Production Society of Kenya (APSK), Nairobi, Kenya, 22-23 March 2000. Available at: http://tinyurl.com/olo2ex6.

Wiggins, S., J. Kirsten and L. Llambí. 2010. The future of small farms. *World Development* 38: 1341-1348.