Full Length Research Paper

Dairy farmer households' farm gate milk price heterogeneity in Kericho County, Kenya

Elijah K. Ng'eno[1,2]

[1]Department of Agricultural Biosystems and Economics, University of Kabianga, Kenya.
[2]Department of Agricultural Economics and Resource Management, Moi University, Kenya.

This study aimed at identifying factors affecting farm gate milk price heterogeneity in dairy farmer households in Kericho County, Kenya. Multistage cluster sampling technique was used to collect data from 432 dairy farmer households. To estimate survey data, multivariate probit and selectivity biased mixed-effects linear regression models were used. Results showed that increased daily milk output sold and number of commercial milk buyers resulted in increased probability of farm gate milk price variability by 3.8 and 12%, respectively. However, number of milking cows and trust levels on commercial milk buyer by seller decreased farm gate milk price heterogeneity by 89 and 87%, respectively. While selling through commercial, milk buyers had significant positive effect on farm gate milk price, majority of dairy farmer households were hesitant to engage with them since milk buyers valued supply security which came from trusted relationships and contracts. Therefore, critical strategies to improve farm gate milk prices are needed. These include strengthening of dairy farmer groups and partnership development, bolstering milk cooperative societies and increased financial investments in livestock milk markets by national and county governments.

Key words: Farm gate milk price, price heterogeneity, mixed-effects linear regression, dairy farmer households, Kericho County, Kenya.

INTRODUCTION

Numerous studies have shown that agricultural markets in developing countries, including sub-Saharan Africa (SSA), are undergoing rapid changes in response to strong economic growth, improved infrastructure and communication systems, and growing demand among consumers for higher quality products. Associated with and facilitating these changes are a range of new interventions and investments, from creative ways to finance value chains, to Information Communication Technology (ICT) solutions for the quick and reliable delivery of market information for farmers, to new organizational approaches for linking small farmers to markets (ILRI, 2011). According to Beneberu et al. (2011), livestock milk marketing is a favorite sector, where African governments choose to intervene in a variety of ways. These interventions range from outright

fixing of wholesale and retail milk prices to monopolizing the export market, yet in many instances policy decisions on livestock milk marketing are taken in the absence of vital information on how they affect small-scale livestock producers, traders and consumers. Very often price fixing at unrealistic levels leads to open black markets, where the real prices substantially differ from those officially listed (Beneberu et al., 2011).

According to FAO (2011), global livestock milk markets offers processors and suppliers increased income and direct cost savings, but they are also posing the threat of market exclusion. A considerable potential to enhance milk market access and marketing success may be available through the promotion of farmer groups, community-based organizations and cooperatives (FAO, 2011). Output prices received by farmers significantly determine their welfare especially in rural areas where there is weak non-farm income which limits diversification of agricultural production amongst producers. While there are debates about the actual and potential impacts of having a wide array of commercial milk buyers on broader welfare of the rural poor; case study evidence suggests that farmers are worst placed when faced with a privately owned or government-controlled monopsony (Gorton and White, 2007; Sadler, 2006). Farmers' welfare depends mostly on the price received for their output in environments of minimal agricultural policy support, the absence of social safety nets, and a weak non-farm rural economy which limits agricultural diversification (Sauer et al., 2012). These features characterize much of Kericho County, where rural poverty is widespread.

In the study county, daily milk price received by dairy farmer households for their milk output had been of considerable concern. Evidence has shown that farm gate milk prices have often been significantly variable and vary considerably between farmer household milk producers. Dairy farmers have been unaware of milk prices received by other farmers due to the weak physical and commercial infrastructure. According to Liefert and Liefert (2007), poor physical and commercial/institutional infrastructure raises transport and transaction costs.

Dairy cooperative societies within Kericho County, which used to be an integral part of the formal milk collection and marketing, have been relegated to buyers of last resort due to their low milk purchase prices. These cooperative societies have been marketing big proportion of their milk directly to processors and urban markets within the county. However, farm gate raw milk prices from the milk buyers have been fluctuating periodically to levels too low to cover farmers' costs of production (KDB, 2015). According to Business Daily (2014), New Kenya Cooperative Creameries (KCC) followed Brookside dairy processor by lowering farm gate milk price and in delaying payment for unsold milk stocks to farmers. New KCC bought raw milk at Kenya Shillings (KES) 32 per litre from KES 40 in March, 2014, while Brookside lowered its farm gate milk price to KES 30 from KES 35

and 40 in March and February, 2014, respectively. By end of March 2014, New KCC had lowered its farm gate milk price payments by 20% while Brookside Dairies had lowered by 14.3 and 25% in March and February, 2014, respectively.

Jari (2009) argues that despite the fact that smallholder farmers face difficulties in marketing, they continue to produce and survive in the face of unfavourable conditions some of which can be solved through use of trusted marketing channels. Farmers maximize return on investments through value addition, complimenting own produce from other sources as well as offering diversified products from the same material inputs. When selling their products, such farmers will use marketing channels that enable their produce to reach the market at least cost per unit of output. By pooling skilled manpower, dairy farmers who are chain actors are able to minimize on transaction costs, access market information and adhere to government regulations more easily.

Dairy farmers are able to take collective action on securing new markets, bargaining for better prices for milk and milk products and use of the most effective marketing channel. Such actions are taken against a background of strong associations by farmers who are trained and have a strong entrepreneurial orientation. However, Sauer et al. (2012) argues that farmers' welfare depends mostly on the price received for their output in environments of minimal agricultural policy support, the absence of social safety nets, and a weak non-farm rural economy which limits agricultural diversification.

Therefore, the motivation of this study was to bridge the literature gap using multivariate probit and selectivity biased mixed-effects linear regression models to explain determinants of price heterogeneity between dairy farmer households and within the county. By considering the causal relationship between participation in selling milk to commercial milk buyers and dairy farmer household welfare, this study anticipated that the findings would address the counterfactual queries that were important in forecasting the impacts of policy changes and that for alleviating dairy farmer household income security in Kericho County, Kenya.

MATERIALS AND METHODS

Sampling procedure

The target population was restricted to 94,427 smallholder livestock milk producers and marketers, divided proportionately amongst the six sub-counties of Kericho County, Kenya. Multistage cluster sampling procedure was then used to get the sample size of interest. The county was clustered into six sub-counties that formed sample sites for the study. To achieve representative sample size, the six sub-counties formed the first-stage cluster that had the target population. These clusters were selected based on the fact that small scale dairy farming was dominant and practiced throughout the six sub-counties. It also reflected significant differences in structure of the dairy milk marketing business in the county. Within the six sub-counties, second-stage cluster sample of

wards and villages with high concentration of small scale dairy farmers was then selected. The sampled milk producing n^{th} smallholder dairy farmer household was determined by the proportionate size sampling methodology (Anderson et al., 2007) as shown in Equation 1.

$$N_0 = \frac{Z^2 pq}{e^2} \tag{1}$$

where N_0 is the sample size, Z is the standard normal value of 1.96 significant at 5% confidence level, e is the margin of error, p is the estimated population proportion of dairy farmers with characteristics of interest, $q = 1-p$, $Z = 1.96$, and e = degree of precision.

Sample units were calculated proportionately based on the number of dairy farmer households in each sub county and as a proportion of the total dairy farmers in the county against the desired sample size of 504. Based on the aforementioned criteria, the random sample of dairy farmer households selling raw milk to different milk marketing channels was set for the whole county consisting of 75 farmers from Kipkelion East, 63 from Kipkelion West, 91 from Kericho West, 44 from Kericho East, 81 from Soin/Sigowet and 150 from Bureti. After data cleaning, 432 observations remained for analysis.

Data

This study used both primary and secondary data. The data was collected through cross-sectional sample design for dairy farmer households. Seasonal observation involved observing the natural behaviour of the dairy farmer households in order to describe the existing situation and to obtain information that were relevant to the goals of the study. Secondary data was obtained from existing published literature desktop literature and internet. Farm records from a sampled of few farmer households were also used to supplement secondary data sources.

Data types

The data types that were used in this study encompassed representative sample of dairy farmer households representing the various household categories, types of commercial and non-commercial milk marketing channels and changing structure of dairy sector. To analyze the responses of milk producers, the study categorized the choice of milk marketing channels into a binary outcome, whether the dairy farmer household sold milk at farm gate to commercial milk marketing channels (Y_1) and if farmer household chose to sell also to final consumers (non-commercial channel) or otherwise (Y_0). Data collected included dairy farmers' socio-economic characteristics, actual milk production, milk market competitiveness and other related obligations with the milk buyers. Farm production data comprised of the size of land under dairy production, average volume of milk produced per day, amount of livestock inputs and farm gate milk prices.

Respondents also provided information regarding market competitiveness and estimated total number of potential commercial buyers for their milk. This would capture the degree of switching power of the dairy farmer household from one commercial buyer to the other. The study also included data on whether the farmer sold per day total milk output on contract signing or on spot cash sale. To capture the trustworthiness of commercial milk buyers, a measure of trust on the commercial milk buyer by the dairy milk farmer household was included. This characteristic was analyzed by a proxy that identified the perception that the dairy milk

farmer had in relation to their trust in the commercial milk buyer.

Regarding milk marketing characteristic, a dummy variable was introduced to capture whether the dairy farmer household sold via milk cooling/chilling plants (cooperative society), milk sheds or through milk bars or not. Time series data on farm gate milk prices received by the farmer household over a period of three years (2013, 2014 and 2015) was also collected from the farmers. This entailed use of pair wise comparison of the six sub county mean milk prices for the three years using Tukey's HSD (honest significant difference) test.

Data analysis and diagnostics

Econometric analysis of data consisted of two stages. Multivariate probit model was used in the first stage to estimate factors which determined milk marketing channel choice decision equation, specifically whether farmers sold raw milk only to a commercial milk buyer or sold also to a final consumer. Secondly, mixed effects linear regression model was used in the analysis of determinants of farm gate milk price heterogeneity in the county. The two empirical models for data analysis were linked by the inverse Mill's ratio (*MR*). The study assumed that it was likely that the characteristics of small scale dairy farmer who sold milk only to a commercial buyer differed from those who sold also to final consumers. However, selection effect or bias correction factor exists in cross sectional data since farmers themselves decide whether or not to sell to a particular marketing channel. Consequently, those who sold and those not could differ systematically, leading to non-random selection bias. Therefore, diagnostic tests were conducted from the regression results of STATA output. To check on multicollinearity, the study used variance inflation factor (VIF) and contingency coefficient (CC) among discrete and continuous variables, respectively. All assumptions were tested and corrected accordingly.

Theoretical framework

According to rational choice theory, individual households' rank mutually exclusive alternative decisions in order of utility and will choose the alternative with maximum expected utility given their socio-economic and demographic characteristics and relevant resource constraints. Hence, in this study, the producer's milk marketing channel choice that fetched better milk price was conceptualized using a random utility model (RUM). It was assumed that economic agents, including smallholder dairy farmers, use certain livestock milk marketing systems only when the perceived utility or net benefit from using such a method is significantly greater than is the case without it. Again, smallholder dairy farmers were assumed to be rational and they want to derive the highest utility from the choices they make; either to market their produce independently or under a certain milk marketing channel depending on the returns. They made their choices with respect to random utility theory. The choice decision maker was guided by unobservable, observable and random characteristics while making a decision. Although utility was not directly observed, the actions of economic agents were observed through the choices they made.

The study formulated milk marketing channel selection/choice decision as a two-alternative choice (selling to commercial milk buyer(s) = 1 and selling to final milk consumer(s) = 0).

Let a decision maker (dairy farmer with raw milk for sale) choose from a set of mutually exclusive alternatives, j = 1, 2,..., J. The decision maker obtains a certain level of utility U_{ij} from each alternative. The discrete choice model is based on the principle that the decision-maker chooses the outcome that maximizes the utility. The producer makes a marginal benefit-marginal cost calculation based on the utility achieved by selling to a market channel or to

another. His/Her utility is not observed, but some attributes of the alternatives as faced by the decision-maker are observed. Hence, the utility is decomposed into deterministic (V_{ij}) and random (ε_{ij}) part:

$$U_{ij} = V_{ij} + \varepsilon_{ij}; \forall_{ij} \in N \tag{2}$$

Since ε_{ij} is not observed, the decision-makers' choice cannot be predicted exactly. Instead, the probability of any particular outcome is derived. The utilities or the difference between benefit and cost cannot be observed directly, but the choice made by the producer reveals which one provides the greater utility (Greene, 2003).
    A producer selects market channel j=1 if;

$$U_{ik} > U_{ij} \forall_j \neq k \tag{3}$$

where U_{ik} denotes a random utility associated with the market channel j = k, and V_{ij} is an index function denoting the producer' average utility associated with this alternative. The second term ε_{ij} denotes a random error which is specific to a producer's utility preference (McFadden, 2000). Now, suppose that Y_i and Y_j represent a household's utility for two milk marketing choices, which are denoted by U_i and U_j, respectively. The linear random utility model, the milk marketing channel choice is modeled as in equation 4.

$$U_{ij} = \beta_j X_{ij} + \varepsilon_{ij} \tag{4}$$

where U_{ij} is a vector of the milk marketing channel choices (j = 1 commercial milk buyers; and 0 for final milk consumers) of ith dairy farmer, β_j is a vector of channel-specific parameters. ε_{ij} is the error term assumed to have a distribution with mean 0 and variance 1 and identically distributed (Greene, 2003). X_{ij} is the vector of explanatory variables that determined and or influenced the perceived desirability of the choice of the milk marketing channel. Therefore, for the case of choice of a livestock milk marketing channel, if a dairy farmer household decides to use option *j* marketing channel, it follows that the perceived utility or benefit from option *j* marketing channel is greater than the utility from other options (say *k*) marketing channel depicted as follows:

$$U_{ij}(\beta_j^1 X_i + \varepsilon_j) > (U_{ik}(\beta_k^1 X_i + \varepsilon_k), \quad k \neq j \tag{5}$$

The probability that a dairy farmer will choose milk marketing channel *j* among the set of livestock milk marketing channels to market his milk instead of the *k* marketing channel could then be defined as:

$$P(Y = 1 \mid X) = P(U_{ij} > U_{ij}) \tag{6}$$

Therefore,

$$P(\beta_j^1 X_i + \varepsilon_j - \beta_k^1 X_i - \varepsilon_k > 0 \mid X) \tag{7}$$

Hence $P(\beta_j^1 X_i - \beta_k^1 X_i + \varepsilon_j - \varepsilon_k > 0 \mid X)$

$$P(X^* X_i + \varepsilon^* > 0 \mid X = F(\beta^* X_i) \tag{8}$$

where *P* is a probability function, U_{ij}, U_{ik}, and X_i are as defined

earlier, $\varepsilon^* = \varepsilon_j - \varepsilon_k$ is a random disturbance term, $\beta_j^* = (\beta_j^1 - \beta_k^1)$ is a vector of unknown parameters that can be interpreted as a net influence of the vector of independent variables influencing the choice of a milk market, and $F(B^* X_i)$ is a cumulative distribution function of the error terms (ε^*) evaluated at $B^* X_i$. The exact distribution of *F* depends on the distribution of the random disturbance term, ε^*. Depending on the assumed distribution that the random disturbance term follows, several qualitative choice models can be estimated (Greene, 2003).
    The choice of a milk marketing channel that offered better milk price was fundamental and important decision for the dairy farmer households. Consistent with the theoretical model, the study assumed that dairy farmer households practiced dairy farming for good milk price and income maximization, to smooth household income through market guarantee and market access and production volume utility respectively despite liquidity constraints. Others practiced dairy farming for prestige. However, liquidity preference played a major role in dairy farmer households' decision for particular milk marketing channel.

Analytical framework

Multivariate probit and mixed effects linear regression models were used in the analysis of the determinants of farm gate milk price heterogeneity in small holder dairy farmer household in Kericho County, Kenya. However, the two empirical models for data analysis were linked by the inverse Mill's ratio (*MR*) inferred in Equation 14. The study assumed that it was likely that the characteristics of small scale dairy farmer household that sold milk only to a commercial buyer differed from those that sold also to final consumers. However, selection effect or bias correction factor exists in cross sectional data since farmers themselves decide whether or not to sell to a particular marketing channel. Consequently, those who sell and those not could differ systematically, leading to non-random selection bias. Estimation and inference problems in econometric models would arise if incorrect non endogeneity assumption about the structure of the decision making process is made. The selection model of Heckman (1979) describes an estimation problem resulting from incomplete data leading to simultaneity problem. Data in the milk marketing channel choice analysis assumed that milk prices would only be observed for the subsample of dairy farmer households who sold milk to commercial buyers. Milk selling depended on covariates that were assumed to affect milk price, or the price that was needed to be offered to induce a dairy farmer to enter the milk market. According to Heckman (1979), econometric model postulates that if a wage offer exceeds the reservation wage, then the wage will be observed for that individual, which was also applied to the milk price in this study. The selection problem was that milk price was only observed for a dairy farmer household that sold milk to a commercial milk buyer and milk price was unobserved or latent for those that sold to final consumers. Therefore, to account for selection bias, two equations were envisaged (adopted from Heckman 1979):

$$Y_0 = X\beta + u_1, \quad \text{(price offered)} \quad \text{and} \quad Y_\tau = X\beta + u_2,$$
(reservation price) $\tag{9}$

It follows then that, $Y = Y_0$ if $Y_0 \geq Y_\tau$ and *Y* is missing or 0 otherwise. Selectivity bias refers to the fact that if this study was to estimate the milk price function based on observations for which there was data, then the estimates of the effects of variables that

determined milk price rates would be inconsistent and biased. Therefore, this study viewed the price equation as the substantive equation of interest and specified as follows:

$$Y = X\beta + u_1, \tag{10}$$

along with a milk selling (participation) equation,

$$P^* = Z\gamma + u_2, \tag{11}$$

in which a dairy farmer household sells milk if $P^* \geq 0,$ and

$$\Pr(P = 1 | Z) = \Phi(Z\gamma) \tag{12}$$

It can be shown that,

$$E(u_1 \mid P = 1) = E(u_1 \mid u_2 > -Z\gamma) = E(u_1 \mid u_2 < Z\gamma) = \sigma_{12}\left[\frac{\phi(Z\gamma)}{\Phi(Z\gamma)}\right] \tag{13}$$

By symmetry of normal distribution, the study had:

$$\frac{\phi(-z)}{1 - \Phi(-z)} = \frac{\phi(z)}{\Phi(z)}$$

and

$$\frac{\phi(-z)}{\Phi(-z)} = \frac{\phi(z)}{1 - \Phi(z)} \tag{14}$$

The ratio $\frac{\phi(z)}{\Phi(z)}$ is the inverse Mills ratio (or non-selection hazard rate) and, $\frac{\phi(z)}{1 - \Phi(z)} = \frac{\phi(z)}{\Phi(-z)}$ is the selection hazard rate. Therefore, from the modeling stand point for this study hypothesis, multi-level modeling approach referred to as mixed-effects or hierarchical model adopted from Bryk and Raudenbush (2002) was used to estimate the determinants of variations in farm gate milk prices for those smallholder farmers who sold milk to commercial milk buyers only in the county as shown in Equation 15.

$$P^* = Z\gamma + u_2, \ P = 1, \text{ if } P^* \geq 0, \text{ 0 otherwise}$$

$$Y^* = X\beta + u_1, \ Y^* = Y \text{ if } P = 1 \tag{15}$$

where P is a probit model; Y^* is a latent continuous dependent variable. In this study, milk prices were treated as continuous. The study also assumed that u_1 and u_2 were distributed as multivariate normal, with means zero, covariances $\sigma_{u_1}^2, \sigma_{u2}^2$, and $\sigma_{u_{12}}^3$. The inverse Mills ratio $\hat{\lambda} = \phi(Z\hat{\gamma})/\Phi(Z\hat{\gamma})$ that was generated in objective one data was then included as a regressor in the mixed effects linear regression analyses model.

Estimation of the model

Mixed-effects linear regression or hierarchical model was used in this study to investigate the determinants of variations in farm gate milk prices for those smallholder farmers who sold to commercial buyers only in the county. The reason was because some of the covariates were grouped according to one or more characteristics. The mixed model was characterized as containing both fixed and random effects. The fixed effects were analogous to standard regression coefficients and were estimated directly. The random effects were not directly estimated but were summarized according to their estimated variances and covariances. According to Sauer et al. (2012) random effects may take the form of either random intercepts or random coefficients, and the grouping structure of the data may consist of multiple levels of nested groups, that is, the error distribution of the linear mixed model is assumed to be Gaussian.

In the study, the dependent variable was farm gate raw milk price in Kenya shillings (KES) per liter received by the dairy farmer household. Milk price data covered three years, with respondents providing average milk price received in the year 2013, 2014, and 2015, respectively. The Laird and Ware (1982) form of the milk price model (Equation 16) was adopted for this study:

$$P_{im} = \alpha + \varepsilon p_{imt-1} + \beta_1 X_{1im} + \beta_2 X_{2ijm} + \beta_3 X_{3ikm} + \beta_4 X_{4ikm} + \beta_5 X_{5ikm} + \beta_6 X_{6ikm} +$$
$$\beta_7 X_{7ikm} + \beta_8 X_{8ikm} \beta_9 X_{9ikm} + \beta_{10} X_{10ikm} + \beta_{11} \lambda_{lm} + \sum_n b_n Z_{1nm} + u_{lm} \tag{16}$$

$$b_n \sim \text{iid } N(0, \gamma b^2), \text{cov}(b_n, b_{n-1}) = \gamma_{n,n-1}, u \sim \text{iid } N(0, \sigma^2 \lambda_{im}), \text{cov}(u_{im} - u_{i-1}, m) = \sigma^2 \lambda_{imi-1}$$
.

The model was estimated by maximum restricted or residual likelihood (REML) (Harville, 1977). Where P_{im} is the price per liter of milk for the farmer i^{th} observation in the mth sub-county; ε, u, μ, ρ, τ, and ϕ are the fixed-effect coefficients which are identical for all the sub-counties m; P_{imt-1}, β_{1im}, β_{2im}, β_{3im}, and β_{4im} are the fixed-effect coefficients which were identical for all the six sub-counties (m); P_{imt-1}, X_{1im}, X_{2im}, X_{3im}, and X_{4im} are the fixed-effect regressors for observation of farmer i in sub-county m (where P_{t-1} is the milk price in 2015; X_1 is the number of milking cows; and X_2 refers to distance to the nearest milk market, X_3 refers to the number of farm gate commercial milk buyers; X_4, X_5, X_6, X_7, X_8 and X_9 are total milk output sold via New KCC, Brookside Ltd, milk traders, self-help groups, home for consumption and milk output sold via milk cooling/chilling plants respectively. X_{10} is a vector of trust-related variable (trust in seller on buyer, a cross effect between trust and percentage of milk output sold to commercial milk buyer); and b_n are the random-effect coefficients for sub-county m, assumed to be multivariate and normally distributed and varying by sub-county; b_n are designed as random variables and are hence similar to the errors u; z_n are the random-effect regressors; γb^2 and $\gamma_{n,n-1}$ are variances and covariances among the random effects assumed to be constant across groups; u_{im} is the error for observation of farmer i in sub-county m assumed to be multivariately normally distributed also assumed to be Gaussian; $\sigma^2 \lambda_{imi-1}$ are the covariances between errors in group m ($\lambda_{imi} = \sigma^2$, $\lambda_{imi-1} = 0$) (observations were sampled independently within the six sub-counties and were assumed to have constant error variance)). Fixed effects were assumed to be similar to standard regression coefficients and were estimated directly. However, the random effects were not directly estimated but were summarized according to their estimated variances and covariances. Random effects took the forms of either random intercepts or random coefficients, and the grouping structure of the data consisted of multiple levels of nested groups related to the six

sub counties and trust levels. X_{11} is the Inverse Mills ratio (λ) obtained from the first stage regression controlling for potential selection bias. The use of this ratio was motivated by the property of the truncated normal distribution (Heckman, 1979).

RESULTS AND DISCUSSION

Characteristics of farmer households

As shown in Table 1 of results, majority of dairy farmer households owned two or one milking cows with the highest being 50 milking cows. The median herd size was 3 milking cows, while 160 of the sampled dairy farmer households owned two milking cows. From the results, 55% of dairy farmer households sold milk to commercial buyer(s), 40% sold to final consumers while 4% sold to both commercial buyer(s) and final consumers. However, some farmers sold milk in more than one market outlet depending on unit price offered, volume of milk produced, and urgency of the need for cash.

Table 2 presents farm gate milk price summary statistics for those dairy farmer households that sold to commercial milk buyers and final consumers. In 2013, 2014 and 2015, the actual average farm gate milk price received by all the farmers selling exclusively to commercial buyers in the county was Kenyan Shillings (KES) 29.91, 32.71 and 35.51 per liter per day, respectively. For the final consumers, the average farm gate milk price for the three years was KES 33.03, 37.31 and 41.65, respectively.

Results in Table 3 show the total number of potential commercial buyers of raw milk at farm gate for the surveyed farmers. Five commercial milk buyers existed in the study area, a clear indication of milk market competitiveness in the county. This also captured the degree of switching power that dairy farmer households had in marketing their raw milk and the degree to which markets were characterized in the county. The results as presented showed the mean milk price offered by the various commercial buyers on farm's milk output sold via milk marketing channels. Milk traders/vendors purchased raw milk at an average farm gate price of KES 36.80 per liter per day from 181 of farmers selling only to commercial buyers in the county. Milk traders/vendors offered a mean minimal price of KES 25 and a maximum price of KES 60 per liter of milk per day. Milk cooling/chilling plants (milk cooperative societies) and milk buying self-help groups bought milk from only 61 (14%) and 8 (2%) of the surveyed dairy farmer households at an average daily farm gate price of KES 33.34 and 37.5 per liter, respectively.

Diagnostic tests

Preliminary results for the diagnostic tests revealed that potential multicollinearity among explanatory variables was found not to have any potential influence on estimates from the model. The highest pair-wise correlation was 0.4, whereas multicollinearity is a serious problem if pair-wise correlation among regressors is in excess of 0.5 (Gujarati, 2004). An analysis of variance inflation factor (VIF) did not show any problem since none of the VIF of a variable exceeded 8 (Greene, 2003). The tests of the fixed effects as presented in Table 4 also provided the F-tests for each of the fixed effects that were specified in the study model. The random effects were not directly estimated but were summarized according to their estimated variances and covariances, that is, represented by random effect and residual covariance. Since the p- value of the F-test for overall significance test was less than the significance level, the null-hypothesis was rejected and conclusion was made that the model used in the study provided a better fit than the intercept-only model.

Heterogeneity in farm gate milk price coefficients

Table 5 presents the results of the mixed-effects linear regression model for the determinants of farm gate milk price in Kericho County obtained through Equation 16. The results also present a summary of the parameters that were used to specify the random effects and residual covariance matrices. Since no repeated effects were specified in the model, the error terms were independent with variance approximately 51.48. The random effects had the scaled identity variance structure, and had a variance parameter, which were approximately 21.7 as shown in the Table 5. The regression was able to explain the variation in the observed farm-gate milk price and hence, most of the coefficients from the regression equation were as expected and the study expectations met. The null hypothesis held that determinants of milk price heterogeneity have no significant effect on farm-gate milk price was thus rejected. The Inverse Mills ratio (IMR) or selectivity bias correlation factor had significant positive effect on the households' farm gate milk price per liter per day.

Results revealed that the number of milking cows was negatively associated with farm-gate milk price for those farmers who sold milk only to commercial milk buyers. As the number of milking cows increased per unit, the dairy farmer household received lower average farm gate milk price per liter per day. Regarding the estimated coefficient of distance, the coefficient negatively influenced farm gate milk price. Distance to milk buying points was associated with decrease in farm-gate milk price. A unit increase in distance to the milk market was associated with 19.20% decrease in farm-gate milk price. The lesser the milk price received by the dairy farmer, the more difficult and costly it would have been to get involved in the milk market. This was because most of the commercial milk buyers were located in trading

Table 1. Milking cows per farm household and by type of milk marketing channel (Authors' Estimates from Survey Data, 2016).

Number of milking cows	HHs selling milk only to commercial buyer(s)	HHs selling milk to final consumers only	HHs selling milk to both final consumers and commercial buyer(s)	Total
1	54	72	4	130
2	104	53	3	160
3	30	32	4	66
4	22	7	6	35
5	12	4	1	17
6 – 9	9	3	2	14
10 – 19	6	2	1	9
Above 20	1	0	0	1
Total	238	173	21	432
Mean = 2.6	SD =3.1	Minimum = 1	Maximum = 50	-

HH: Household.

Table 2. Milk price statistics for farmer households selling to both commercial buyers and final consumers (Authors' Estimates from Survey Data, 2016**).**

All counties	Mean selling to commercial	Standard deviation	N	Mean selling to final consumers	Standard deviation
Average milk price received 2015 (KES)	35.51	6.93	194	41.65	8.44
Average milk price received 2014 (KES)	32.71	5.88	194	37.31	8.01
Average milk price actually 2013 (KES)	29.91	5.55	194	33.03	7.03

Table 3. Commercial farm gate milk buyers (Authors' Estimates from Survey Data, 2016).

Commercial farm gate milk buyer	No. of HHs	Mean price per liter	Standard deviation	Minimum	Maximum
New KCC	9	34.78	7.10	30	50
Brookside	14	33.93	7.32	27	50
Traders/Vendors	181	36.76	7.49	25	60
Milk buying Self-help groups	8	37.5	5.35	30	45
Milk cooling/chilling plants	61	33.34	6.67	26	60
Total	273	-	-	-	-

Table 4. Tests of fixed effects (Authors' Estimates from Survey Data, 2016).

Source	Numerator Df	Denominator df	F	Sig.
Number of farm gate commercial buyers	1	417.000	0.217	0.041
Number of milking cows	1	417.000	46.847	0.000
Distance to milk market	1	417.000	4.343	0.038
Total milk output sold via New KCC	1	417.000	0.003	0.057
Total milk output sold via Brookside	1	417.000	1.437	0.031
Total milk output sold via milk Traders	1	417.000	8.254	0.004
Total milk output sold via Self-help groups	1	417.000	0.209	0.048
Total milk output home consumption	1	417.000	5.556	0.019
Total milk output sold via milk cooling/chilling plants	1	417.000	0.338	0.561
IM Ratio	1	417.000	146.326	0.000
Trust	4	417.000	2.543	0.039
Intercept	1	0.000	40.011	0.000

Table 5. Determinants of farm gate milk price.

Parameter	Coefficients	Standard error	T	P>|t|
Number of milking cows	-0.889	0.1298	-6.845	0.000*
Distance to milk market	-0.192	0.0920	-2.084	0.038**
Number of farm gate commercial buyers	0.120	0.2982	-0.466	0.041**
Total milk output sold via New KCC	0.002	0.0360	0.054	0.957
Total milk output sold via Brookside	-0.032	0.0266	-1.199	0.231
Total milk output sold via milk Traders	0.038	0.0132	-2.873	0.004*
Total milk output sold via Self-help groups	0.019	0.0412	0.457	0.648
Total milk output home consumption	-0.077	0.0324	-2.357	0.019**
Total milk output sold via milk cooling/chilling plants	0.0095	0.0164	0.581	0.541
Trust	-0.878	0.363	-2.42	0.016**
Probability of random selection				
Inverse mills ratio (λ)	17.354	1.4346	12.097	0.000
Intercept	29.491	4.9598	5.946	0.000
Random effect and residual covariance				
Residual (variance/covariance)	51.4813	3.5653	-	0.000*
Intercept variance	21.7407	-	-	-
LR test vs. linear regression chi^2(4)	9.88	-	-	-
Log- restricted-likelihood	-1513.75	-	-	-
Prob > chi^2	0.043	-	-	-

Trust levels (5 Point Likert scale) – 1 = strongly disagree, 2 = Disagree, 3 = Undecided, 4 = Agree, 5 = strongly agree. * = 1%, ** = 5%, and *** = 10% levels of significance. *p*-value of Likelihood Ratio Test (Pr>x^2). Source: Authors' Analytical Computation from Survey Data, 2016

centers and yet the majority dairy farmer households were in the villages. Ultimately, this may have become a limiting factor for farmers from such areas to sell more proportions of milk to commercial milk buyers. In addition, milk being a perishable product, dairy farmers feared the risk of losing their milk during long distance transportation in addition to high transport costs involved in formal milk marketing channels. This finding was in convergence with findings of Muricho (2015), who observed that as the distance increases away from the farm, there is a decline in the transacted quantities by farmers.

Output milk price received by dairy farmer households was also of considerable importance to this study. The number of farm gate commercial milk buyers' exerted positive and significant effect on farm gate milk price. Results showed that a unit increase in the number of commercial milk buyers was associated with a 12% increase in farm gate milk price per day. As expected, farmers that produced large volumes of milk receive better milk price for their milk supply because of the existing competition between the milk buyers. According to Swinnen and Maertens (2007), greater competition should lead to more equal rent sharing, evidenced by higher producer prices and more services for farmers. Farmers would receive less milk price when faced with a privately owned or government controlled monopoly as argued by Sadler and Good (2006). An improvement in milk price was expected to significantly lead to an

increase in household income. This would have stimulated demand for varied household goods that would eventually lead to improvement of the welfare of the dairy farmer household. However, the number of commercial milk buyers may not have been an effective measure of competition particularly where there was collusion amongst the buyers.

Selling milk through milk traders as a determining factor presented a statistically significance and positive result on farm gate milk price per liter per day for farmer households selling milk to commercial milk buyers. The model predicted that an addition of one trader to the milk market would lead to an increase by 3.8% in the farm gate milk price per liter per day. This meant that the relationship level between commercial milk buying traders and dairy milk selling farmer households had positive significant affect on farm gate milk price in the long run. The sign of the coefficient was consistent with the expectations, that is, the higher the relationship between milk buying traders and dairy milk seller, the more the farmer would prefer the milk traders in the long run. Thus, the result was consistent with the hypothesis, indicating a preference for direct sales to milk traders by the dairy farmer households in return for better milk price. In the context of milk market liberalization and its consequences and off-farm/non-farm income opportunities being limited to most of the dairy farmers in Kericho County, positive association of selling milk

through milk traders with farm gate milk price per liter per day was expected. Since liberation in Kenya, milk buying traders have become diversified. The sheer number of milk traders has led to buying and selling of milk to consumers in raw unprocessed or unpackaged form due to the consumers' unwillingness to pay the extra costs of processing and packaging. As a consequence, large numbers of milk traders enhances competition for milk supply, thereby increasing the family's daily income from milk sales. These milk markets may also provide valuable opportunities for rural and urban employment. However, unprocessed and unpackaged milk is prone to diseases which may be hazardous to the final consumers. Product quality was significantly linked to higher farm gate milk price.

Results on trust revealed a negative relationship with the actual milk price received by the dairy farmers. Although the coefficient of the factor trust showed negative sign (reject hypothesis), which would accept the formulated hypotheses on trust level, it was statistically significant, that is, it did have significant impact on farm gate milk price variability. The negative coefficient on trust by the smallholder dairy farmers on commercial milk buyer proved to be aligned with the expectations for farm gate milk sales price. A smallholder dairy farmer household saw the commercial milk buyer as the greatest source of knowledge about the farm gate milk price and understood that farm gate milk price provided by the commercial milk buyer may have been better in solving potential problems. When trust levels by dairy farmer on commercial milk buyers increased, the commercial buyer would eventually have no interest in him or her. However, given the statistical significance of the coefficient, this study concluded that the effects of trust interfered with farm gate milk price.

The Inverse Mills ratio (λ), which was a correction factor for selectivity bias, was significant and depicted that there were unobserved factors that would have affected the selection (milk marketing channel choice price) as well as the outcome (marketed milk surplus) equation.

CONCLUSIONS AND POLICY IMPLICATIONS

This paper analyzed the determinants of milk price heterogeneity in dairy farmer household farms in Kericho County, Kenya using multivariate probit and selectivity bias mixed-effects linear regression model. Post-hoc pair wise comparisons on mean milk price heterogeneity at farm-gate received in all the six sub-counties was partially supported. Mean milk price received by dairy farmer households in 2013 was significantly better than the mean milk price received in 2014. Based on the selectivity bias mixed-effects linear regression model, results revealed that the number of milking cows, distance to milk market and trust on commercial milk buyer by seller were statistically significant. However, the

factors were associated with decrease in farm-gate milk price per liter per day. Farm gate commercial milk buyers and percent total milk output sold via milk traders were the main determinants of variability in farm gate milk price per liter per day for the farmer households that sold milk to commercial buyers only. The other remaining factors were not identified statistically as determinants of variability in farm gate milk price per liter per day. In conclusion, the present study contributes to our theoretical understanding by showing that the development of factor relationship characteristics may influence the choice decision of a smallholder dairy farmer milk seller, but a commercial milk marketing channel choice may not always be the first option for the dairy milk farmer household. The welfare of the smallholder dairy farmer therefore depends on the milk price received for their daily milk output. Besides being a valuable source of income for rural dairy farmer households in Kericho County, dairy milk production also helps in smoothing household incomes, which in turn smoothens consumption hence improving farmer household welfare over long periods of time. Financial, market access and market involvement seems to be very important factors affecting farm gate milk price per day. Thus, local county government and national governments could pay more attention to enhance dairy farmer household access to milk markets and financial investment. In future relevant stakeholders should redesign or reform milk marketing implementation strategies or improve/strengthen existing milk marketing policy.

CONFLICT OF INTERESTS

The author has not declared any conflict of interests.

REFERENCES

Anderson DR, Sweeny JD, Williams TA, Freeman J, Shoesmith E (2007). Statistics for business and economics. 9th Edition. Thomson Learning. Thomson Press (India) Ltd.

Beneberu TK, Girma AH (2011). Determinants of Sheep Prices in the Highlands of Northeastern Ethiopia: Implication for Sheep Value Chain Development. Addis Ababa, Ethiopia. http://doi.org/10.1007/s11250-011-9837-x.

Bryk AS, Raudenbush SW (2002). Hierarchical Linear Models: Applications and Data Analysis Methods (Advanced Quantitative Techniques in the Social Sciences) (2nd Ed.). Stanford: Sage Publications.

Business Daily (2014). Kenya Dairy Board plans to regulate raw milk prices, p. 12. Nairobi. Retrieved from http://www.businessdailyafrica.com/Corporate-News/Kenya.

FAO (2011). World Livestock 2011: Livestock in Food Security World. (A. McLeod, Ed.). Rome, Italy: FAO.

Gorton M, White J (2007). Transformation and Contracting in the Supply Chains of the Former Soviet Union: Evidence from Armenia, Georgia, Moldova, Ukraine and Russia. Global Supply Chains, Standards and the Poor", Wallingford, Oxfordshire: CABI Publishing.

Greene WH (2003). Econometric Analysis (5th ed.). Upper Saddle River, NJ.: Prentice Hall, Upper Saddle River, NJ.

Gujarati DN (2004). Basic Econometrics. 4th Edition, New York: Tata

McGraw Hill Publisher.

Harville DA (1977). Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems. Am. Stat. Assoc. 72:320-338.

Heckman JJ (1979). Sample Selection Bias as a Specification Error. Econometrica (Vol. 47). The Econometric Society. Retrieved from http://www.jstor.org/stable/1912352.

ILRI (2011). Towards priority actions for market development for African farmers. In Proceedings of an international conference, Nairobi, Kenya,13-15 May 2009. Nairobi, Kenya. P 400. Retrieved from http://mahider.ilri.org/handle/10568/16491.

Jari B (2009). Institutional and technical factors influencing agricultural marketing channel choices amongst smallholder and emerging farmers in the Kat River Valley. Afr. J. Agric. Res. 4(11):1129-1137.

Kenya Dairy Board (KDB) (2015). The Kenyan Dairy Industry at a Glance. http://www.kdb.co.ke/

Laird M, Ware JH (1982). Random-effects models for longitudinal data. J. Biomet. 38(38):963-974.

Liefert W, Liefert O (2007). Distortions to Agricultural Incentives in Russia. Agricultural Distortions Working Paper 08, (No. 08). Washington, DC.

McFadden D (2000). Economic Choices. Economic Science. Berkeley, USA.

Muricho G, Kassie M, Obare G (2015). Determinants of Market Participation Regimes among Smallholder Maize Producers in Kenya International Maize and Wheat Improvement Center, Kenya (No. 12).

Sadler M (2006). Comparative analysis of cotton supply chains in Central Asia. In: Swinnen, J.F.M. (Ed.), Case Studies on Vertical Co-ordination in Agrofood Chains in Europe and Central Asia, ECSSD. Washington, DC: World Bank.

Sauer J, Gorton M, White J (2012). Marketing, Cooperatives and Price Heterogeneity: Evidence from the CIS Dairy Sector. Agric. Econ. 43(2):165-177.

Swinnen MM, Maertens JF (2007). Trade, Standards and Poverty: Evidence from Senegal. Oxford, UK.