Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Networks, cooperatives and the milk market: evidence on outlet choice from urban and peri-urban dairy farms in Ethiopia

Tesfaye Solomon Eshetu

tesyeshi@gmail.com

Ethiopian Institute of Agricultural Research

Jema Haji

Haramaya University

Dawit Alemu

Adam Bekele

Ethiopian Institute of Agricultural Research

Research Article

Keywords: milk market channels, multivariate probit, supply network, egocentric network analysis, cooperatives, urban and peri-urban dairy, Ethiopia

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11798301/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Networks, cooperatives and the milk market: evidence on outlet choice from urban and peri-urban dairy farms in Ethiopia

Abstract

Raw milk from urban and peri-urban dairy farms reaches Ethiopian consumers through direct sales, retailers and wholesalers, yet little is known about how the structure of a farm's trading network shapes the outlets it uses. We combine a farm survey with egocentric network data on suppliers and buyers collected from dairy farms in Addis Ababa, Bishoftu, Sebeta and Mekelle (157 interviews, 133 with complete data) and estimate a multivariate probit model of sales to consumers (38% of farms), retailers (63%) and wholesalers (13%). One farm in five used more than one outlet, and consumer and retailer sales were strong substitutes ($\rho = -0.87$). Cooperative members were 19 percentage points less likely to sell directly to consumers and 16 and 18 points more likely to sell to retailers and wholesalers. A one-standard-deviation increase in ego-network density was associated with a 20-point lower probability of selling to consumers and a 12-point higher probability of selling to wholesalers; network interconnectedness, and accessibility when entered alone, were negatively associated with selling to retailers. These associations were stable across alternative specifications, whereas the positive association of density with retail sales was not. Education was associated with direct sales. The findings suggest that closed trading circles may steer farmers away from direct sales and that cooperatives channel milk towards intermediaries, which makes the relationships through which farmers trade worth considering in dairy marketing policy.

Keywords: milk market channels; multivariate probit; supply network; egocentric network analysis; cooperatives; urban and peri-urban dairy; Ethiopia

JEL codes: Q13; D85; O12

1 Introduction

Dairy production in Ethiopia is mostly smallholder-based: in 2019 about 6.7 million dairy cows produced an estimated 3.6 billion litres of milk, over 95% of it from local breeds (CSA 2019). Urban and peri-urban dairy farming supplies fresh milk to growing cities and provides income and employment in East Africa (Gillah et al. 2012). In Ethiopia, which has one of the largest cattle populations in the world but remains a net importer of dairy products (as summarised by Mamo et al. 2021), most of this milk is marketed informally. In a survey of 475 commercial dairy farms in nine cities and towns, 71% of farms sold milk directly to consumers, whereas only about 9% supplied formal processors, and cooperatives received a very small share of the milk sold (Mamo et al. 2021). The outlet a farm uses matters: in the same study, direct sales to consumers earned higher net prices than sales to traders and

processors, and prices and demand are depressed in fasting seasons. The choice of outlet therefore affects revenue, the regularity of sales and the effort spent on marketing.

A sizeable literature has examined the determinants of milk outlet choice in Ethiopia. Most studies relate the choice to household and farm characteristics, access to information and extension, cooperative membership, distance and prices. Multinomial logit studies in Wolaita and central Ethiopia find that cooperative membership, market information and extension raise the probability of selling through cooperative outlets and that herd size and prices shape the use of processors and hotels (Kuma et al. 2013; Mengistu and Meressa 2023). Studies using multivariate probit models, which allow farms to use several outlets and allow unobserved factors to be correlated across outlets, report that education, experience, farm size, distance, cooperative membership, prices and location affect outlet choice among commercial urban and peri-urban farms (Mamo et al. 2021). Almost all of this work treats the farmer as an isolated decision-maker.

Economic sociology and the study of agricultural trade suggest that this is incomplete. Farmers' choices are embedded in relationships with suppliers, buyers and peers that carry information, trust and obligations (Granovetter 1985; Uzzi 1996). Evidence from Ethiopia and elsewhere shows that such relationships can shape how farmers sell. Ethnic and religious ties influence whether Ethiopian potato growers trade through middlemen, and personalised relationships may lock farmers into intermediated trade even when this lowers income (Abebe et al. 2016). In India, farmers' position in their social network and the number of their connections are associated with having long-term relationships with traders (Songsermsawas et al. 2015). Social capital raises the returns of traders in Africa (Fafchamps and Minten 2002), and peers affect technology adoption (Bandiera and Rasul 2006). In agricultural economics, however, networks have been studied mainly as channels of information for technology adoption (Bandiera and Rasul 2006; Conley and Udry 2010; Maertens and Barrett 2013), including dairy technology in Ethiopia (Amlaku et al. 2012). We are not aware of studies that relate the measured structure of a dairy farm's egocentric network to the choice among milk outlets.

This paper builds on our earlier analysis of the same farms, which showed that the structure of a farm's supply network is associated with dairy technology adoption: denser networks were associated with lower adoption and more interconnected networks with higher adoption (Solomon et al. 2022). Here we ask whether the same structural characteristics are associated with a different outcome, the outlets through which farms sell their milk.

This paper fills the gap using survey and egocentric network data from urban and peri-urban dairy farms in four Ethiopian milk-sheds. It asks how three structural characteristics of a farm's supply network (density, interconnectedness and accessibility) are associated with the probability of selling raw milk to consumers, retailers and wholesalers, after controlling for farmer and farm characteristics and institutional links. The paper makes three contributions. First, it brings measured network structure into

the analysis of milk outlet choice. Second, it uses a multivariate probit that recognises that outlets can be combined, tests whether outlet decisions are interdependent, and reports unconditional and conditional marginal effects. Third, it checks the robustness of the network results to collinearity among the network measures, to site effects and to the exclusion of the thinly observed wholesale outlet.

The remainder of the paper is organised as follows. Section 2 gives the context and the conceptual framework. Section 3 presents the data and the empirical model. Section 4 reports the results, Section 5 discusses them and their limitations, and Section 6 concludes.

2 Context and conceptual framework

2.1 Milk marketing in urban and peri-urban Ethiopia

Ethiopia's major milk-sheds, including Addis Ababa and its surroundings and Mekelle, differ in the development of processing and collection systems (Brandsma et al. 2012). Around Addis Ababa, farms can sell to processors and collection centres as well as to households and hotels, whereas in regional cities such as Mekelle direct sales to consumers and hotels dominate and collection centres are scarce (Mamo et al. 2021). Farms typically sell to three broad types of buyer. Consumers buy at the farm gate or through door-to-door delivery and pay the highest prices, but the farmer bears delivery effort and demand risk. Retailers, including shops, tea and coffee houses and small milk vendors, buy smaller lots to resell. Wholesalers and collectors buy larger volumes for onward sale to retailers, hotels and processors. Cooperatives play a limited role in raw milk marketing in these areas (Mamo et al. 2021).

2.2 Transaction costs, embeddedness and network structure

Transaction cost economics explains outlet choice by the costs of searching for buyers, negotiating, monitoring quality and enforcing payment, which are high for a perishable product without cooling (Williamson 1985; Holloway et al. 2000). Relationships can lower these costs, but the effect depends on their structure. We adopt the supply network perspective, which looks beyond the buyer–seller pair to the web of suppliers, service providers and customers around a focal firm (Kim et al. 2011; Basole et al. 2018). Following Bellamy et al. (2014), the structural characteristics of this network, such as how effectively the firm can reach information through it and how its partners are linked to one another, determine the flows of information and knowledge. We argue that they also shape the flow of trading opportunities. We consider three characteristics of a farm's ego network, that is, the farm and its first-tier suppliers and buyers.

Density is the share of possible ties among the farm's partners that exist. In a dense network, partners know and monitor one another, which supports trust and repeated, norm-governed exchange (Coleman 1988; Uzzi 1996). Such closure lowers the cost of regular trade with known intermediaries, but it also limits access to new customers, which is the function of bridging ties (Granovetter 1973; Burt 1992). The effect of density is contested: closure can support trust and collaboration (Ahuja 2000; Carnovale

and Yeniyurt 2015), whereas sparse networks give access to diverse, non-redundant information and opportunities (Burt 1992, 2001), and embeddedness in strong ties can lock firms into existing relationships (Grabher 1993). Which effect dominates depends on the context. Interconnectedness captures how redundant the farm's contacts are. It is defined as one minus Burt's network efficiency, so that a high value means that the farm's partners are largely connected to one another and offer little non-overlapping reach (Burt 1992). Accessibility is the farm's information centrality, that is, how quickly information can flow between the farm and all its partners through all paths (Stephenson and Zelen 1989).

The theory gives clear expectations for some variables and not for others. We expect dense networks to be associated with intermediated outlets (retailers and wholesalers) and with fewer direct sales to consumers, because closure favours repeated trade with known buyers while direct sales need reach beyond the closed group. Redundant contacts (high interconnectedness) offer fewer independent buyers, but we have no outlet-specific prediction, and the same holds for accessibility. Among the controls, cooperatives pool milk and bargain collectively, so members should sell larger lots to intermediaries and less directly to consumers (Holloway et al. 2000; Bernard and Spielman 2009). Education and labour affect marketing capacity and the feasibility of labour-intensive direct delivery, but the sign for each outlet is open. Because theory predicts competing effects for several variables, we treat the analysis as exploratory and report all outlets and all measures, not only those with expected signs. Figure 1 summarises the framework.

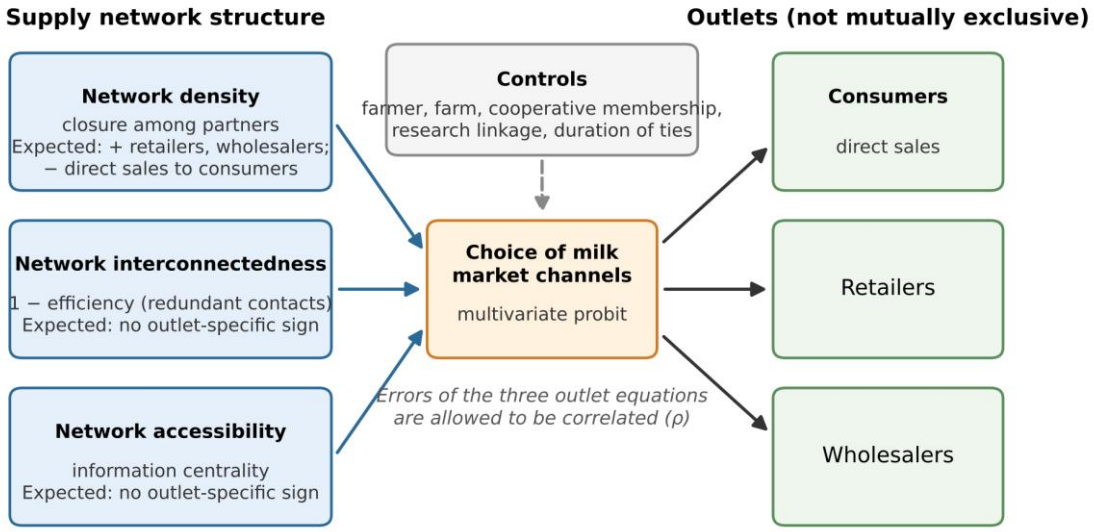


Figure 1. *Conceptual framework: supply network structure and the choice of milk market channels. Source: author's conceptualisation.*

3 Data and methods

3.1 Study areas, sampling and data collection

The data come from the baseline survey of the Ethiopia Control of Bovine Tuberculosis Strategies (ETHICOBOTS) project, a five-year research project of a consortium of epidemiologists, geneticists, immunologists and social scientists in Ethiopia and the United Kingdom. The study sites were selected from the project's baseline results to capture the patterns of dairy supply networks across locations: Addis Ababa, Bishoftu and Sebeta in the Addis Ababa milk-shed, and Mekelle in Tigray (Figure 2). A multi-stage sampling design was used. The study area was stratified into three strata by proximity to market infrastructure (near, middle and away from markets), one sub-city or town was selected in each stratum, a kebele or woreda was chosen in each on the basis of dairy production, and 169 dairy farms were randomly selected from the chosen areas.

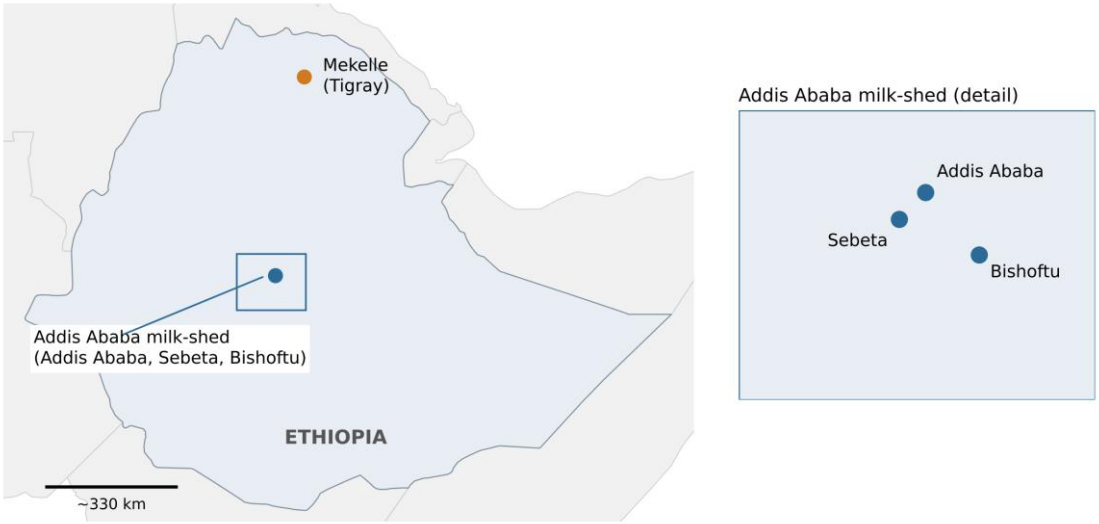


Figure 2. *Location of the study sites. Source: author's drawing from Natural Earth boundaries and survey locations.*

Data were collected by trained enumerators using computer-assisted personal interviewing programmed in CSPro, with a supervisor present and after a pre-test on farms similar to the target population. Participants were informed of the voluntary and confidential nature of the study and gave consent before the interview. Two instruments were used. A structured questionnaire recorded farm and farmer characteristics, and a second instrument programmed in EgoNet recorded the ego network of each farm, based on its links with suppliers and buyers in terms of goods exchange during the survey period: the ego–alter and alter–alter ties and the attributes of the alters. After listwise deletion of missing covariates, including the network measures, 133 farms are used in the estimation (18 in Sebeta, 44 in Mekelle and 71 in Addis Ababa and Bishoftu).

3.2 Dependent variables

Three binary variables record whether the farm sold raw milk to (i) consumers, (ii) retailers or (iii) wholesalers (1 = yes). They are not mutually exclusive. Sales to farm-gate buyers, processors and supermarkets are recorded in the data but are not modelled due to few observations.

3.3 Network and control variables

Three structural measures were computed from each farm's ego network, following the definitions used in our earlier analysis (Solomon et al. 2022). All ties are binary and undirected.

Accessibility is measured by information centrality (Stephenson and Zelen 1989), $\mathrm{IC}_i = [C_{ii} + (\Sigma_j\, C_{jj} - 2\Sigma_j\, C_{ij})/n]^{-1}$, where **C** is the inverse of the matrix $\mathbf{B} = \mathbf{D} - \mathbf{A} + \mathbf{J}$, **A** is the adjacency matrix of the ego network, **D** the diagonal matrix of the number of direct ties of each node, **J** a matrix of ones and n the number of nodes. The index has a minimum of zero and no upper bound; it is higher the more and the shorter the paths that connect the farm to other nodes.

Interconnectedness is one minus Burt's network efficiency, $\mathrm{Inter}_i = 1 - \mathrm{Eff}_i$, with $\mathrm{Eff}_i = \Sigma_j\, [1 - \Sigma_q\, p_{iq}\, m_{jq}] \,/\, n_i$, where p_{iq} is the share of the farm's ties invested in partner q, m_{jq} equals 1 if partners j and q are tied and 0 otherwise, and n_i is the number of partners (Burt 1992; Wasserman and Faust 1994). A high value means that the farm's partners are largely connected to one another, so that many ties are redundant.

Density is the number of ties observed among the farm's N alters divided by the number of possible ties, $N(N-1)/2$ (Borgatti 1997; Carnovale and Yeniyurt 2015). It is scaled from 0 to 1 here.

The three measures are mean-centred before estimation. Control variables describe the farmer (sex, education), the farm (age, labour, sales and herd size category), institutional links (cooperative membership, contact with support institutions) and the relationship with partners (years of interaction). Following Solomon et al. (2022), small farms are those with five or fewer cattle, medium farms have more than five cattle, and large farms are licensed state or private commercial farms; the model includes a dummy for small farms. Table 1 gives definitions and summary statistics for the estimation sample.

Table 1. Explanatory variables and summary statistics (N = 133)

Variable	Definition	Mean	SD	Min	Max
Sex (male)	1 = male respondent	0.61	0.49	0	1
Farm age	Years since establishment	11.10	8.87	1	32
Education	Years of schooling of the owner or manager	7.57	4.70	0	16
Labour	Square root of the number of hired employees	0.89	0.66	0	3.16
Cooperative member	1 = member of a dairy producers' cooperative	0.41	0.49	0	1
Sales	Log of annual farm sales revenue in 2018/19, birr (winsorised)	11.50	0.73	9.39	12.52
Interaction	Mean duration of relationships with supply partners (years)	4.76	5.15	0	30
Small farm	1 = five or fewer cattle	0.41	0.49	0	1

Variable	Definition	Mean	SD	Min	Max
Research linkage	Number of support institutions contacted (research institutes, agricultural offices, universities) [TO CONFIRM]	0.51	1.13	0	5
Network density	Ego-network density, mean-centred	0.00	0.20	−0.37	0.49
Network accessibility	Information centrality, mean-centred	−0.01	0.79	−0.91	3.05
Network interconnectedness	1 − efficiency, mean-centred	0.00	0.21	−0.23	0.58

3.4 Empirical model

Multinomial models assume that a farmer picks exactly one outlet. Farms can sell through several outlets, and unobserved factors such as managerial ability or risk attitudes can affect several decisions, so we use a multivariate probit (Greene 2012), which is a pick-any model (Walsh 1995; Kim et al. 2007; Aurier and Mejia 2014). For farm i and outlet m (consumers, retailers, wholesalers), the latent propensity is

$$y^*_{im} = \mathbf{x}_i'\boldsymbol{\beta}_m + \varepsilon_{im}, \text{ and } y_{im} = 1 \text{ if } y^*_{im} > 0 \text{ and 0 otherwise,}$$

where $\mathbf{x}_i$ includes the network measures and controls and the errors ε are jointly standard normal with correlation matrix $\mathbf{R}$. The off-diagonal elements ρ measure the correlation between unobserved determinants of two outlet decisions; a negative value indicates substitutes and a positive value complements. If all ρ equal zero, separate probits would suffice, which we test with a likelihood ratio test.

The coefficients $\boldsymbol{\beta}$ are not effects on probabilities. We report three average marginal effects. The unconditional effect is the change in $P(y_{im} = 1) = \Phi(\mathbf{x}_i'\boldsymbol{\beta}_m)$ when a variable changes. The conditional effect is the change in $P(y_{im} = 1 \mid \text{the other outlets} = c)$, where $c = 0$ (none of the other two outlets is used, cond_0) or $c = 1$ (both are used, cond_1). This conditional probability is the joint probability of the outcome pattern divided by the probability of the other outcomes, computed from the trivariate normal distribution (Sodjinou and Henningsen 2012). For 0/1 variables we report the discrete change from 0 to 1 and for continuous variables the derivative per unit, averaged over the sample. For the network measures we also report the effect of a one-standard-deviation increase, because one unit lies outside their range. Standard errors use the delta method.

3.5 Estimation, replication and robustness

The model was estimated by simulated maximum likelihood with the Geweke–Hajivassiliou–Keane (GHK) simulator (Geweke 1989; Börsch-Supan and Hajivassiliou 1993; Keane 1994) using the mvProbit package in R with 5,000 draws (Henningsen 2012). The results were replicated independently in Python with a GHK simulator and scrambled Halton draws; coefficients agree with the R estimates to within simulation error, and the code is available as a Jupyter notebook. The GHK procedure is described by Gates (2006) and was implemented for Stata by Cappellari and Jenkins (2003). We check robustness in four ways: (i) entering each network measure separately, because the measures are highly correlated; (ii) dropping interconnectedness; (iii) adding site dummies in a bivariate model of the two

common outlets, because only 17 farms sell to wholesalers; and (iv) comparing conditional with unconditional effects.

4 Results

4.1 Outlet use and farm characteristics

Retailers were the most common outlet (63% of the 133 farms, 84 farms), followed by consumers (38%, 51 farms) and wholesalers (13%, 17 farms) (Figure 3a). Most farms used a single outlet (retailers only 44%, consumers only 26%, wholesalers only 4%), 19.5% combined outlets, and 6% used none of the three (Table 2). The most frequent combinations were consumers with retailers and retailers with wholesalers; consumers and wholesalers were rarely combined. Outlet use differed by site (Figure 3b). Direct sales to consumers were reported by 20% of farms in Mekelle, compared with 46% in Addis Ababa and Bishoftu and 50% in Sebeta, whereas retail sales were similar (66%, 63% and 56%) and wholesale sales ranged from 7% in Mekelle to 17% in Addis Ababa and Bishoftu.

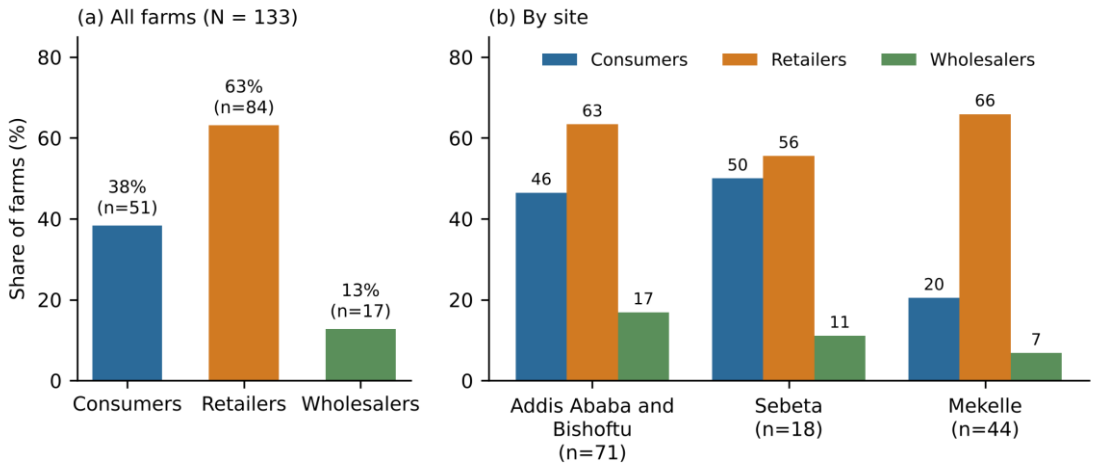


Figure 3. *Outlet use among 133 dairy farms: (a) share of farms using each outlet; (b) share by site. Farms can use more than one outlet, so shares sum to more than 100%. Source: author's survey.*

Table 2. Combinations of outlets used (N = 133)

Combination	Farms	Share (%)
Retailers only	59	44.4
Consumers only	35	26.3
Consumers and retailers	14	10.5
Retailers and wholesalers	10	7.5
Wholesalers only	5	3.8
Consumers and wholesalers	1	0.8
All three	1	0.8
None of the three	8	6.0

Figure 4 and Table 3 compare farms that use and do not use each outlet. The largest raw differences are for cooperative membership and network structure. Cooperative members make up 82% of farms selling

to wholesalers but 35% of the others, and 28% of farms selling to consumers but 49% of the others. Farms selling to consumers have lower network density, and farms selling to retailers have lower interconnectedness and accessibility. These bivariate patterns motivate the multivariate analysis, which holds the other variables constant.

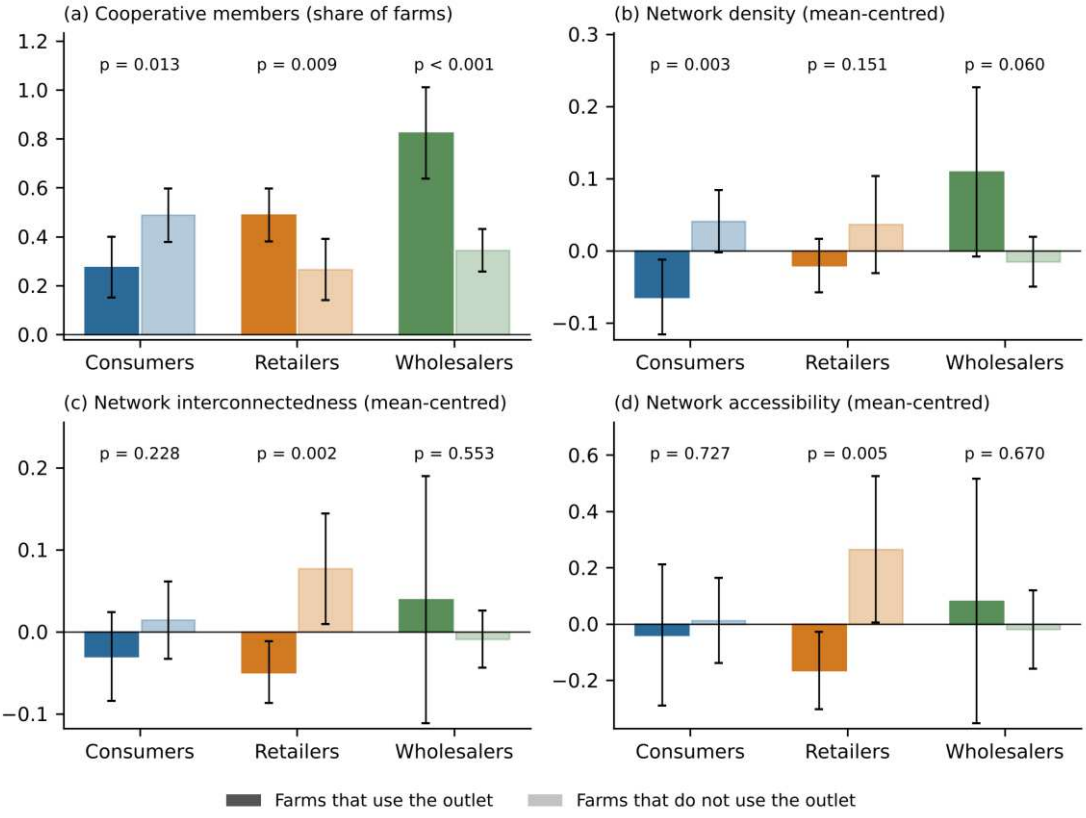


Figure 4. *Cooperative membership and network structure of farms that use and do not use each outlet (means with 95% confidence intervals; p-values from Welch t-tests). Network measures are mean-centred. Source: author's survey.*

Table 3. Mean characteristics of farms that use and do not use each outlet

Variable	Consumers: use	Consumers: non-use	Retailers: use	Retailers: non-use	Wholesalers: use	Wholesalers: non-use
Cooperative member (share)	0.28	0.49 (p = 0.013)	0.49	0.27 (p = 0.009)	0.82	0.35 (p < 0.001)
Education (grades)	8.59	6.94 (p = 0.046)	7.14	8.31 (p = 0.175)	7.35	7.60 (p = 0.846)
Labour (square root)	0.97	0.84 (p = 0.245)	0.82	1.00 (p = 0.112)	0.99	0.87 (p = 0.392)
Research linkage	0.37	0.60 (p = 0.244)	0.48	0.57 (p = 0.645)	0.53	0.51 (p = 0.935)
Network density	−0.064	0.041 (p = 0.003)	−0.021	0.036 (p = 0.151)	0.109	−0.016 (p = 0.060)
Network interconnectedness	−0.030	0.014 (p = 0.228)	−0.049	0.077 (p = 0.002)	0.039	−0.009 (p = 0.553)
Network accessibility	−0.040	0.013 (p = 0.727)	−0.166	0.265 (p = 0.005)	0.081	−0.020 (p = 0.670)

Notes: p-values are from Welch two-sample t-tests of equal means between users and non-users of each outlet.

The three network measures are closely related (Figure 5). Density correlates with interconnectedness at 0.79, interconnectedness with accessibility at 0.89 and density with accessibility at 0.64, and the relationships are close to linear for most farms. This strong overlap matters for the interpretation of the regression results, and we return to it in Section 4.7.

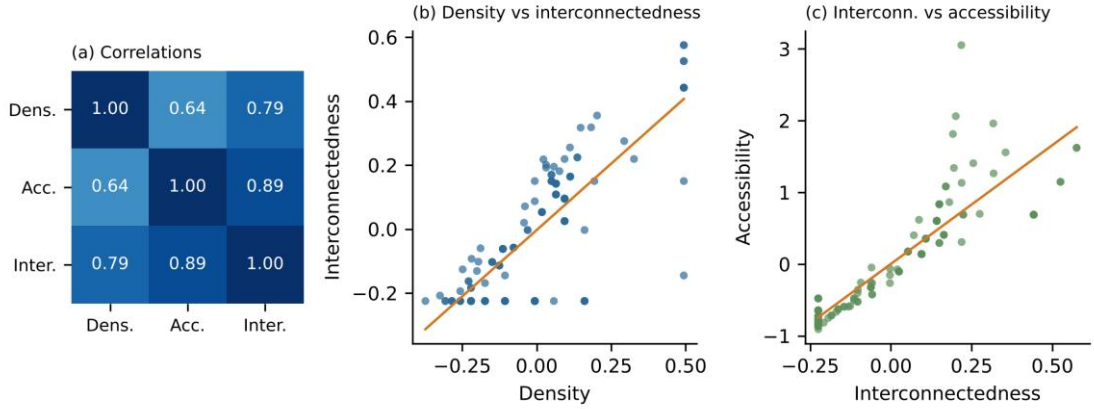


Figure 5. *Relationships among the three network measures: (a) correlation matrix; (b) density and interconnectedness; (c) interconnectedness and accessibility (N = 133; fitted lines from simple linear regressions). Source: author's survey.*

4.2 Model fit and interdependence of outlet choices

The likelihood ratio test rejects the independence of the three equations ($\chi^2(3) = 43.46$, $p < 0.001$). The errors of the consumer and retailer equations are strongly and negatively correlated ($\rho = -0.87$, SE 0.07), whereas the wholesaler equation is uncorrelated with the other two ($\rho = -0.07$, SE 0.25, and 0.04, SE 0.25). Unobserved factors that raise the probability of selling directly to consumers lower the probability of selling to retailers, so the two outlets are substitutes. The correlation is much larger than the −0.33 found between traders and consumers among commercial dairy farms in a wider sample (Mamo et al. 2021), which may reflect the narrower outlet definitions used here.

4.3 Determinants of outlet choice

Table 4. Multivariate probit coefficients (standard errors in parentheses)

Variable	Consumers	Retailers	Wholesalers
Sex (male)	−0.204 (0.264)	0.356 (0.263)	0.177 (0.366)
Farm age	0.018 (0.017)	−0.007 (0.017)	−0.035 (0.031)
Education	0.068** (0.030)	−0.039 (0.029)	−0.047 (0.041)
Labour	0.317 (0.231)	−0.443* (0.240)	0.086 (0.346)
Cooperative member	−0.585** (0.266)	0.503* (0.263)	1.062*** (0.364)
Sales (log)	−0.071 (0.181)	0.023 (0.185)	−0.154 (0.245)
Interaction	−0.041 (0.030)	0.026 (0.028)	0.050 (0.038)
Small farm	0.439 (0.319)	−0.338 (0.318)	−0.249 (0.453)
Research linkage	−0.255** (0.130)	0.065 (0.114)	0.093 (0.184)
Network density	−3.576*** (1.268)	2.405* (1.295)	2.895** (1.261)
Network accessibility	0.129 (0.432)	0.340 (0.365)	0.250 (0.533)
Network interconnectedness	1.667 (2.004)	−4.928** (1.940)	−2.223 (2.082)
Constant	−0.038 (2.097)	0.448 (2.169)	0.237 (2.806)

Notes: ***, **, * denote significance at 1%, 5% and 10%. N = 133; log-likelihood = −163.83; ρ(consumers, retailers) = −0.868*** (0.069); ρ(consumers, wholesalers) = −0.072 (0.246); ρ(retailers, wholesalers) = 0.041 (0.252). Source: author's estimates.

4.4 Unconditional marginal effects

Table 5 translates the coefficients into effects on the probability of using each outlet. Cooperative members were 18.5 percentage points less likely to sell to consumers ($p < 0.05$) and 16.0 and 18.2 points more likely to sell to retailers ($p < 0.05$) and wholesalers ($p < 0.01$). Each additional grade of education raised the probability of direct sales by 2.1 points ($p < 0.05$), and each unit of research linkage lowered it by 7.9 points ($p < 0.05$). More labour was associated with a 13.8-point lower probability of selling to retailers (per unit of the square-root measure, $p < 0.10$). Sex, farm age, sales, length of interaction and small herd size were not significant for any outlet.

Table 5. Unconditional average marginal effects on the probability of using each outlet (probability units)

Variable	Consumers	Retailers	Wholesalers
Sex (male)	−0.064 (0.083)	0.113 (0.083)	0.028 (0.057)
Farm age	0.006 (0.005)	−0.002 (0.005)	−0.006 (0.005)
Education	0.021** (0.009)	−0.012 (0.009)	−0.008 (0.006)
Labour	0.099 (0.071)	−0.138* (0.071)	0.014 (0.056)
Cooperative member	−0.185** (0.081)	0.160** (0.081)	0.182*** (0.063)
Sales (log)	−0.022 (0.056)	0.007 (0.058)	−0.025 (0.040)
Interaction	−0.013 (0.009)	0.008 (0.009)	0.008 (0.006)
Small farm	0.135 (0.095)	−0.105 (0.097)	−0.039 (0.069)
Research linkage	−0.079** (0.038)	0.020 (0.035)	0.015 (0.030)
Network density	−1.114*** (0.358)	0.749* (0.389)	0.469** (0.196)
Network accessibility	0.040 (0.135)	0.106 (0.112)	0.040 (0.086)
Network interconnectedness	0.519 (0.618)	−1.535*** (0.560)	−0.360 (0.334)

Notes: effects are per unit of each variable (discrete change for 0/1 variables); delta-method standard errors in parentheses. The network measures have standard deviations of about 0.2 (density, interconnectedness) and 0.8 (accessibility), so per-unit effects are large; see Table 7.

4.5 Conditional marginal effects

Table 6 reports effects conditional on the farm using none of the other two outlets (cond_0) or both of them (cond_1). The model reproduces the observed conditional shares well (for example 80% predicted against 81% observed for consumer sales among farms using neither other outlet). The conditional effects keep the signs of the unconditional effects for cooperatives, education, density and interconnectedness but are mostly imprecise. Only interconnectedness on retailer sales (cond_0: −1.01 per unit; cond_1: −1.61; both $p < 0.05$) and research linkage on consumer sales (cond_0: −0.07, $p < 0.10$) are significant. This is expected, because the cond_1 scenario occurs for only 2–15 farms, so we treat these estimates as illustrative.

273

274

275

276

Table 6. Conditional average marginal effects (probability units; standard errors in parentheses)

Variable	Cons. cond_0	Cons. cond_1	Retail. cond_0	Retail. cond_1	Whole. cond_0	Whole. cond_1
Sex (male)	0.012 (0.076)	0.001 (0.063)	0.063 (0.060)	0.092 (0.090)	0.029 (0.069)	0.025 (0.052)
Farm age	0.005 (0.005)	0.004 (0.005)	0.001 (0.004)	0.002 (0.006)	−0.006 (0.006)	−0.005 (0.006)
Education	0.014 (0.009)	0.012 (0.010)	0.000 (0.006)	0.002 (0.011)	−0.007 (0.008)	−0.006 (0.008)
Labour	0.007 (0.067)	0.020 (0.054)	−0.066 (0.051)	−0.097 (0.089)	0.017 (0.061)	0.014 (0.053)
Cooperative member	−0.083 (0.085)	−0.071 (0.095)	0.043 (0.055)	0.051 (0.111)	0.186 (0.122)	0.155 (0.129)
Sales (log)	−0.020 (0.055)	−0.018 (0.044)	−0.004 (0.041)	−0.010 (0.069)	−0.027 (0.045)	−0.023 (0.037)
Interaction	−0.008 (0.009)	−0.007 (0.009)	0.001 (0.006)	0.000 (0.010)	0.008 (0.007)	0.007 (0.008)
Small farm	0.070 (0.088)	0.068 (0.090)	−0.021 (0.070)	−0.021 (0.112)	−0.037 (0.074)	−0.031 (0.069)
Research linkage	−0.073* (0.039)	−0.060 (0.042)	−0.021 (0.025)	−0.043 (0.045)	0.012 (0.033)	0.010 (0.033)
Network density	−0.667 (0.415)	−0.587 (0.482)	0.087 (0.281)	0.021 (0.498)	0.455 (0.289)	0.386 (0.376)
Network accessibility	0.125 (0.129)	0.092 (0.111)	0.105 (0.083)	0.180 (0.148)	0.047 (0.111)	0.041 (0.080)
Network interconnectedness	−0.577 (0.604)	−0.330 (0.519)	−1.009** (0.436)	−1.610** (0.763)	−0.383 (0.602)	−0.332 (0.349)

Notes: Cons. = consumers, Retail. = retailers, Whole. = wholesalers. cond_0 = the other two outlets are not used; cond_1 = both other outlets are used. Effects are per unit (discrete change for 0/1 variables); delta-method standard errors; **, * denote significance at 5% and 10%.

4.6 Network structure

Network density had opposite signs for direct and intermediated outlets. A one-standard-deviation increase in density (0.20) was associated with a 19.7-point lower probability of selling to consumers ($p < 0.01$) and 13.8- and 11.8-point higher probabilities of selling to retailers ($p < 0.05$) and wholesalers ($p < 0.05$) (Table 7). Interconnectedness lowered the probability of selling to retailers by 32.8 points per standard deviation ($p < 0.01$), with no significant effect on the other outlets. Accessibility had no significant effect in the joint model.

Table 7. Change in the probability of using each outlet for a one-standard-deviation increase in network measures (probability units)

Network measure (SD)	Effect type	Consumers	Retailers	Wholesalers
Density (0.20)	Unconditional	−0.197*** (0.053)	0.139** (0.065)	0.118** (0.057)
Density (0.20)	cond_0	−0.154 (0.107)	0.025 (0.049)	0.113 (0.073)
Density (0.20)	cond_1	−0.090 (0.072)	−0.006 (0.106)	0.098 (0.096)
Interconnectedness (0.21)	Unconditional	0.112 (0.135)	−0.328*** (0.105)	−0.061 (0.045)
Interconnectedness (0.21)	cond_0	−0.099 (0.145)	−0.308** (0.154)	−0.064 (0.083)
Interconnectedness (0.21)	cond_1	−0.081 (0.079)	−0.233* (0.120)	−0.055 (0.045)
Accessibility (0.79)	Unconditional	0.032 (0.108)	0.080 (0.081)	0.035 (0.079)
Accessibility (0.79)	cond_0	0.086 (0.071)	0.067 (0.041)	0.041 (0.104)
Accessibility (0.79)	cond_1	0.077 (0.103)	0.148 (0.122)	0.035 (0.076)

Notes: discrete change for a +1 SD change; cond_0 and cond_1 as in Table 6; delta-method standard errors in parentheses; ***, **, * denote significance at 1%, 5% and 10%.

4.7 Robustness

The three network measures are strongly correlated in the estimation sample (Figure 5), and the variance inflation factor of interconnectedness is about 8.5, so Table 8 shows the network coefficients when the measures are entered separately, when interconnectedness is dropped and when site dummies are added in a bivariate consumer–retailer model. The association of density with lower consumer sales is stable in every specification, and its positive association with wholesale sales is stable wherever the wholesale equation is estimated. The positive association of density with retail sales is not stable: it is significant only at 10% in the joint model and is small or negative when density enters alone or with accessibility. In contrast, interconnectedness and accessibility are negatively associated with retail sales when entered separately (−1.84, $p < 0.01$, and −0.45, $p < 0.01$) and in the model without interconnectedness (accessibility −0.49, $p < 0.05$), even though accessibility is insignificant in the full joint model, where it is almost collinear with interconnectedness. The cooperative coefficients are stable across all specifications: negative for consumers (−0.52 to −0.70), positive for retailers (+0.43 to +0.57) and positive for wholesalers (+1.03 to +1.11, $p < 0.01$). Adding site dummies in the bivariate model leaves the network results unchanged. A 16-variable specification with site dummies, medium herd size and AI services could not be estimated reliably for wholesalers because of complete separation with only 17 users, and we do not report it.

Table 8. Network coefficients across specifications

Specification (N)	Network measure	Consumers	Retailers	Wholesalers
Joint model (133)	Density	−3.58***	2.41*	2.89**
Joint model (133)	Interconnectedness	1.67	−4.93**	−2.22
Joint model (133)	Accessibility	0.13	0.34	0.25
Density only (135)	Density	−1.73***	−0.94	1.68**
Interconnectedness only (135)	Interconnectedness	−0.59	−1.84***	0.59
Accessibility only (135)	Accessibility	−0.07	−0.45***	0.15
Without interconnectedness (133)	Density	−2.77***	0.30	2.21**
Without interconnectedness (133)	Accessibility	0.39	−0.49**	−0.25
Bivariate with site dummies (133)	Density	−3.64***	2.45*	—
Bivariate with site dummies (133)	Interconnectedness	2.16	−5.06***	—
Bivariate with site dummies (133)	Accessibility	0.05	0.34	—

Notes: multivariate probit coefficients on mean-centred network measures with the controls of Table 4; ***, **, * denote significance at 1%, 5% and 10%. Standard errors: density only 0.67, 0.61, 0.83; interconnectedness only 0.61, 0.62, 0.74; accessibility only 0.16, 0.17, 0.22; without interconnectedness (density) 0.94, 0.86, 1.10 and (accessibility) 0.24, 0.24, 0.32; bivariate with site dummies (density) 1.28, 1.29, (interconnectedness) 2.02, 1.95, (accessibility) 0.44, 0.37.

4.8 Summary against the theoretical expectations

Table 9. Summary of results

Variable	Expectation (Section 2.2)	Result
Network density	Positive for retailers and wholesalers, negative for consumers	Consumers: negative, robust. Wholesalers: positive, robust. Retailers: positive in the joint model only, not robust
Interconnectedness	No outlet-specific prediction	Retailers: negative, robust. Others: not significant
Accessibility	No outlet-specific prediction	Retailers: negative when entered alone; not significant in the joint model
Cooperative membership	Negative for consumers, positive for retailers and wholesalers	Confirmed for all three outlets in all specifications
Education	No clear sign	Positive for direct sales to consumers
Labour	No clear sign	Negative for retailers (10% level)
Research linkage	No clear sign	Negative for direct sales to consumers

5 Discussion

5.1 Main findings

The retailer–consumer correlation is the most striking result. Although the multivariate probit allows joint use, only one farm in five combined outlets, and $\rho = -0.87$ shows that the two main outlets behave as alternatives. This supports the joint estimation and tells us that the decision is largely a choice between selling to intermediaries and selling directly to consumers. It also reinforces the finding that consumers and traders are substitutes in the milk marketing of Ethiopian commercial farms (Mamo et al. 2021).

Cooperatives. Members were less likely to sell directly to consumers and more likely to sell to retailers and wholesalers, in every specification. Mamo et al. (2021) also found that cooperative membership lowers the probability of direct sales to consumers among commercial farms. Our results are consistent with cooperatives pooling members' milk for sale in larger lots to intermediaries, which reduces the effort of door-to-door selling and the transaction costs of finding buyers (Holloway et al. 2000). Studies of Ethiopian cooperatives also show that their reach and benefits are uneven: poorer farmers tend not to

join and may be excluded from decisions (Bernard and Spielman 2009), so membership may proxy for farm scale and resources.

Education, labour and research linkage. Education was associated with direct sales to consumers, which fits the argument that selling directly demands marketing and negotiation skills, although the data do not test this and education may proxy for urban consumer contacts. Mamo et al. (2021) report that education raises the probability of selling to processors and hotels, so the role of education may depend on which outlets are available. Labour was associated with fewer retail sales only at the 10% level. The negative association of research linkage with direct sales is hard to interpret and may reflect production systems or location, and we draw no conclusion from it before the variable is clarified.

Network structure. Dense ego networks, where the farm's partners know one another, were associated with intermediated outlets (wholesalers robustly, retailers weakly) and against direct consumer sales. One reading, consistent with embeddedness theory, is that closed circles of repeated partners lower the cost of regular trade with known intermediaries but limit access to new customers (Uzzi 1996; Burt 1992). It also echoes evidence that personalised, group-based relationships can lock Ethiopian smallholders into trading through intermediaries (Abebe et al. 2016). Interconnectedness and accessibility were negatively associated with retail sales, and the measures are so collinear that we cannot separate them. They capture how embedded the farm is among partners who are themselves connected, for example groups of suppliers or neighbouring farms, and farms in such positions may rely less on retail buyers, but our data do not allow us to say why. In India, a more central position in the village network was associated with long-term relationships with traders (Songsermsawas et al. 2015), which suggests that the direction of network effects depends on the type of ties and the market. Our earlier analysis of the same farms found that density was negatively and interconnectedness positively associated with the adoption of dairy technology, which we read as evidence of lock-in in a low-dynamic industry (Solomon et al. 2022; Grabher 1993). The present results are consistent with that reading: farms in dense networks are less likely to sell directly to consumers, an outlet that requires reaching beyond the existing circle of partners. The sign of interconnectedness differs between the two outcomes, positive for adoption and negative for retail sales. A possible explanation, which we cannot test, is that redundant contacts help the circulation of technical knowledge among suppliers but do not help the search for buyers. We therefore stress the robust results, the negative association of density with consumer sales and the positive association with wholesale sales, and treat the retail results as suggestive.

5.2 Implications

The results point tentatively in three directions. First, because cooperative membership steers farms towards intermediaries, strengthening cooperatives' own marketing and their links with processors and milk collection centres could help move milk from informal to formal channels, which Mamo et al. (2021) identify as a priority. Second, farms that wish to expand direct sales may benefit from training

in marketing and from ties beyond their close circle of partners, such as new customers reached through groups or digital platforms; this is a hypothesis and not a tested intervention. Third, the very low share of direct sales in Mekelle suggests that policies need to be tailored to the milk-shed, since in regional cities collection infrastructure is scarcer (Mamo et al. 2021). We do not claim that any outlet is more profitable, because prices and margins by outlet were not analysed.

5.3 Limitations

First, the data are cross-sectional and the networks include buyers as well as suppliers, so the outlet used may shape the network as much as the network shapes the outlet; all results are associations. In our adoption analysis, a nonlinear GMM test with instruments (sex, education, farm age, duration of ties and degree centrality) did not reject the exogeneity of the network measures (Solomon et al. 2022), but that test concerned a different outcome and most of those variables are controls in the present model, so we cannot rely on it here. Second, the sample is small, wholesalers are few (17 farms) and listwise deletion reduced the 157 completed interviews to 133, which limits power and the number of covariates. Third, important transaction cost variables such as distance to market, price information, volume marketed and cooling facilities are not in the model, and omitted variables may be correlated with network structure. Fourth, only three of the recorded outlets are modelled, and the conditional effects given use of the other two outlets describe a scenario seen in only 2–15 farms. Fifth, the three network measures are highly correlated, which makes their separate effects hard to identify. Sixth, the estimates come from four urban and peri-urban sites and may not generalise to rural areas.

5.4 Future research

Panel data, supplier-only network measures or a control-function approach with excluded instruments for network position would help to address reverse causality. Adding prices, margins and transaction cost variables would allow outlet choice to be linked to farm income. Qualitative work could clarify what the ties between farms, cooperatives and intermediaries consist of. Finally, replication in other milk-sheds would test whether the network associations hold where collection centres and processors are more common.

6 Conclusions

This study shows that the structure of a dairy farm’s supply network is associated with where it sells raw milk. Farms with dense networks are less likely to sell directly to consumers and more likely to sell to wholesalers, farms whose partners are highly interconnected are less likely to sell to retailers, and cooperative members shift from direct sales towards retailers and wholesalers. Consumer and retailer sales are strong substitutes. These associations are robust to collinearity among network measures and to the inclusion of site effects, apart from the weaker evidence on density and retail sales. Because the data are cross-sectional and the wholesale outlet is thinly observed, the results describe associations and

motivate further work with panel data and richer transaction cost measures. They suggest that interventions aimed at milk marketing should consider the relationships through which farmers trade as well as the infrastructure and prices they face.

Declarations

Funding: This work was funded by the Biotechnology and Biological Sciences Research Council, the Department for International Development, the Economic and Social Research Council, the Medical Research Council, the Natural Environment Research Council and the Defence Science and Technology Laboratory under the Zoonoses and Emerging Livestock Systems (ZELS) programme, ref. BB/L018977/1, through the Ethiopia Control of Bovine Tuberculosis Strategies (ETHICOBOTS) project.

Competing interests: The author declares no competing interests.

Ethics approval:

This study collected data through focus group discussions (FGDs) and key informant interviews (KIIs) with livestock market experts and cattle traders between January and March 2019. It was a socio-economic survey involving no clinical, biomedical, or otherwise invasive procedures. All participants were informed of the purpose of the study and gave verbal informed consent before taking part; no personal or identifying information about individual participants was recorded, and participation was voluntary. This study was not submitted to a separate ethics committee or institutional review board. Instead, ethical clearance for this study, conducted as part of the Ethiopia Control of Bovine Tuberculosis Strategies (ETHICOBOT) project under the Zoonoses and Emerging Livestock Systems (ZELS) programme (ref: BB/L018977/1), was reviewed by the Agricultural Economics Research Directorate and approved by the Office of the Director General, Ethiopian Institute of Agricultural Research (EIAR) (Ref. No. 21.1/2398/2026, 12 October 2026).

All methods were carried out in accordance with relevant institutional and national guidelines and regulations.

Consent to participate: Participants were informed of the purpose of the study, that participation was voluntary and confidential, and gave consent before the interview.

Consent for publication: Not applicable.

Data availability:

The datasets generated and analysed during the current study (market-level cattle trade links, origin-destination flows, and centrality measures) are available from the corresponding author on reasonable request.

Code availability: The Python replication code (Jupyter notebook) is available.

Author contributions:

TS conceptualization and methodology; TS collected and curated the data; TS analyzed and wrote the original draft; TS, JH, DA and AB wrote the review and edited the manuscript. All the authors read and approved the final version of the manuscript.

Acknowledgements: We thank the members of the ETHICOBOTS consortium, the Ethiopian Institute of Agricultural Research and Kulumsa Agricultural Research Center, the enumerators and the participating farmers.

Author details

Tesfaye Solomon[1,*], Jema Haji[1], Dawit Alemu[2] and Adam Bekele[3]

[1]Haramaya University, College of Agricultural and Environmental Sciences, School of Agricultural Economics and Agribusiness, Dire Dawa, Ethiopia

[2]Stichting Wageningen Research Ethiopia (SWR), Addis Ababa, Ethiopia

[3]Ethiopian Institute of Agricultural Research Institute, Addis Ababa, Ethiopia

References

Abebe GK, Bijman J, Royer A (2016) Are middlemen facilitators or barriers to improve smallholders' welfare in rural economies? Empirical evidence from Ethiopia. J Rural Stud 43:203–213

Ahuja G (2000) Collaboration networks, structural holes, and innovation: a longitudinal study. Admin Sci Q 45(3):425–455

Amlaku A, Sölkner J, Puskur R, Wurzinger M (2012) The impact of social networks on dairy technology adoption: evidence from Northwest Ethiopia. Int J AgriScience 2(11):1062–1083

Aurier P, Mejia V (2014) Multivariate logit and probit models for simultaneous purchases: presentation, uses, appeal, and limitations. Rech Appl Mark (English Edition) 29(2):75–94

Bandiera O, Rasul I (2006) Social networks and technology adoption in northern Mozambique. Econ J 116(514):869–902

Basole RC, Ghosh S, Hora MS (2018) Supply network structure and firm performance: evidence from the electronics industry. IEEE Trans Eng Manag 65(1):141–154

Bellamy MA, Ghosh S, Hora M (2014) The influence of supply network structure on firm innovation. J Oper Manag 32:357–373

Bernard T, Spielman DJ (2009) Reaching the rural poor through rural producer organizations? A study of agricultural marketing cooperatives in Ethiopia. Food Policy 34(1):60–69. https://doi.org/10.1016/j.foodpol.2008.08.001

Borgatti SP (1997) Structural holes: unpacking Burt's redundancy measures. Connections 20(1):35–38

Brandsma W, Mengistu D, Kassa B, Yohannes M, Lee J (2012) The major Ethiopian milksheds. Livestock Research Report 735. Wageningen UR Livestock Research, Wageningen

Burt RS (1992) Structural holes: the social structure of competition. Harvard University Press, Cambridge, MA

Burt RS (2001) Structural holes versus network closure as social capital. In: Lin N, Cook KS, Burt RS (eds) Social capital: theory and research. Aldine de Gruyter, New York, pp 31–56

Börsch-Supan, A., & Hajivassiliou, V. A. (1993). Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models. *Journal of econometrics*, *58*(3), 347-368.

Cappellari L, Jenkins SP (2003) Multivariate probit regression using simulated maximum likelihood. Stata J 3(3):278–294

Carnovale S, Yeniyurt S (2015) The role of ego network structure in facilitating ego network innovations. J Supply Chain Manag 51(2):22–46

Coleman JS (1988) Social capital in the creation of human capital. Am J Sociol 94:S95–S120

Conley TG, Udry CR (2010) Learning about a new technology: pineapple in Ghana. Am Econ Rev 100(1):35–69

CSA (2019) Agricultural sample survey 2018/2019: livestock and livestock characteristics. Central Statistical Agency, Addis Ababa, Ethiopia.

Fafchamps, M., & Minten, B. (2002). Returns to social network capital among traders. *Oxford economic papers*, *54*(2), 173-206.

Gates R (2006) A Mata Geweke–Hajivassiliou–Keane multivariate normal simulator. Stata J 6(2):190–213

Geweke, J. (1989). Bayesian inference in econometric models using Monte Carlo integration. *Econometrica: Journal of the Econometric Society*, 1317-1339.

Gillah KA, Kifaro GC, Madsen J (2012) Urban and peri urban dairy farming in East Africa: a review on production levels, constraints and opportunities. Livest Res Rural Dev 24(11)

Grabher G (1993) The weakness of strong ties: the lock-in of regional development in the Ruhr area. In: Grabher G (ed) The embedded firm: on the socioeconomics of industrial networks. Routledge, London, pp 255–277

Granovetter M (1985) Economic action and social structure: the problem of embeddedness. Am J Sociol 91(3):481–510

Granovetter MS (1973) The strength of weak ties. Am J Sociol 78(6):1360–1380

Greene WH (2012) Econometric analysis, 7th edn. Pearson Education, Harlow

Henningsen, A. (2012). 'mvProbit': Multivariate Probit Models. R package version 0.1-0.

Available from http://CRAN.R-project.org/package=mvProbit

Holloway G, Nicholson C, Delgado C, Staal S, Ehui S (2000) Agroindustrialization through institutional innovation: transaction costs, cooperatives and milk-market development in the east-African highlands. Agric Econ 23(3):279–288

Keane, M. P. (1994). A computationally practical simulation estimator for panel data. *Econometrica*, 62(1):95-116.

Kim J, Allenby GM, Rossi PE (2007) Product attributes and models of multiple discreteness. J Econom 138(1):208–230

Kim Y, Choi TY, Yan T, Dooley K (2011) Structural investigation of supply networks: a social network analysis approach. J Oper Manag 29(3):194–211

Kothari CR (2004) Research methodology: methods and techniques, 2nd edn. New Age International, New Delhi

Kuma B, Baker D, Getnet K, Kassa B (2013) Factors affecting milk market outlet choices in Wolaita zone, Ethiopia. Afr J Agric Res 8(21):2493–2501

Maertens A, Barrett CB (2013) Measuring social networks' effects on agricultural technology adoption. Am J Agric Econ 95(2):353–359

Mamo T, Haji J, Bekele A, Teklewold T, Berg S, Moore HL, Hodge C (2021) Determinants of milk marketing channel selection by urban and peri-urban commercial dairy producers in Ethiopia. Ethiop J Agric Sci 31(3):69–100

Mengistu, A. T., & Meressa, A. M. (2023). Dairy farmers' choice of milk market outlets: evidence from farm households in central Ethiopia. *Journal of Food Quality*, *2023*(1), 5684470.

Sodjinou E, Henningsen A (2012) Community-based management and interrelations between different technology adoption decisions: innovations in village poultry farming in Western Africa. FOI Working Paper No. 2012/11

528 Solomon T, Haji J, Alemu D, Bekele A (2022) The influence of supply network on dairy technology
529 adoption: empirical evidence from urban and peri-urban dairy farming systems of Ethiopia.
530 Ethiop J Agric Sci 32(4):30–55

531 Songsermsawas T, Baylis K, Chhatre A, Michelson H, Prasanna S (2015) Friends or traders? Do
532 social networks explain the use of market mechanisms by farmers in India. Paper presented at the
533 29th International Conference of Agricultural Economists, Milan, 9–14 August 2015

534 Stephenson K, Zelen M (1989) Rethinking centrality: methods and examples. Soc Netw 11(1):1–37

535 Uzzi B (1996) The sources and consequences of embeddedness for the economic performance of
536 organizations: the network effect. Am Sociol Rev 61(4):674–698

537 Walsh JW (1995) Flexibility in consumer purchasing for uncertain future tastes. Mark Sci 14(2):148–
538 165

539 Wasserman S, Faust K (1994) Social network analysis: methods and applications. Cambridge
540 University Press, Cambridge

541 Williamson OE (1985) The economic institutions of capitalism. Free Press, New York