Explorative analysis of relationships among breed, growth traits, prices and sex of sheep using structural equation modeling

Iyiola-Tunji, A. O., Baba, D. and Buba, W.

National Agricultural Extension and Research Liaison Services,
Ahmadu Bello University, Zaria, Nigeria

Corresponding author: tunjiyiola@naerls.gov.ng +2347034234136

Abstract

The dynamics of factors that cumulate to determine the prices of sheep during festive periods (Eid al-Adha) are known. However, such factors that dictate prices during off-season (non-festivity periods) are not fully known. Structural equation modeling as a multivariate analytical tool is capable of bringing out the latent (hidden) factors responsible for price determination for sheep. Prices, body weight and linear body measurements of 621 sheep of both sexes consisting of Balami, Sudanese, Uda and Yankasa breeds were used to examine relationships among the variables. The sheep were obtained from Unguwa Uku, Dorayi, Kara, Bachirawa, Yankaba, Mariri and Kabara livestock markets in Municipal, Taruauni, Gmale, Nassarawa, Fagge and Dala Local Government Areas of Kano State. Body weights, neck length, back length, leg length, height at wither, loin girth and chest girth were parameters monitored. The data obtained were subjected to generalized least squares (GLS) estimator of a triangular seemingly unrelated regression (SUR) model. Maximum likelihood estimation method of structural equation modelling was used to generate parameter estimates through recursive system with correlated errors (SEM Command Language). Price of sheep is an observed variable while type and body parameters were generated as latent variables. Type as a latent variable had direct relationships with breed, sex and body weight class. Body parameter on the other hand had direct relationships with body weight and all the linear body measurements. The z-values were 2.9 for breed, 2.8 for sex and 2.5 for body weight class, others are 4.6 for birth weight, 1.9 for neck length, 4.1 for back length, 6.8 for loin, 7.4 for chest girth, 8.3 for height at wither and 6.1 for leg length. Simultaneous prediction equations for estimating prices of sheep had been generated for some selected livestock markets in Kano State.

Keywords: Sheep; structural equation modelling; price of sheep; path diagram

Introduction

Path analysis is a subset of Structural Equation Modeling (SEM), the multivariate procedure. The model allows examination of a set of relationships between one or more independent variables, either continuous or discrete, and one or more dependent variables, either continuous or discrete. SEM deals with measured and latent variables. A measured variable is a variable that can be observed directly and is measurable. Measured variables are also known as observed variables, indicators or manifest variables. A latent variable is a

variable that cannot be observed directly and must be inferred from measured variables. Latent variables are implied by the covariance among two or more measured variables. They are also known as factors (i.e., factor analysis), constructs or unobserved variables. SEM is a combination of multiple regression and factor analysis. Path analysis deals only with measured variables. Because of the problems associated with pricing of sheep, there is a need for a quick estimate of price of sheep in an open market. This will encourage sheep farmers to raise their stock

purposely for market at a reduced time for profit (Iyiola-Tunji *et al.*, 2011). There are many complex inter-relationships among dependent and independent variables in sheep and its pricing mechanisms in the market. Previous attempt by Iyiola-Tunji *et al.* (2011) used the traditional model generating technique (regression analysis) to examine the inter-relationships between body measurements and prices of sheep. Many of the limitations of traditional statistical models can be understood by recognizing their inflexible architecture and inability to represent networks of causal relations (Grace, 2006). Some more techniques for examining multiple relationships (simultaneously) among productive traits of livestock had been explored. Yakubu *et al.* (2011) used path analysis to explore relationships between conformation traits and milk yield of Bunaji cows. The path analysis is a multiple regression that examined the associations among many independent and dependent variables. It is not however capable of generating latent variables (constructs of factors that are not directly observed or measured) from the set of variables being considered. Structural equation modelling therefore was developed to incorporate both latent and observed variables

(Schumaker and Lomax, 2010). This study was therefore aimed at exploring the interrelationships among the latent and observed variables of breeds, growth traits, sex and prices of sheep using structural equation modelling.

Materials and methods

Prices, body weight and linear body measurements of 621 sheep of both sexes consisting of Balami, Sudanese, Uda and Yankasa breeds were used to exploratively examine relationships among the variables. The sheep were obtained from Unguwa Uku, Dorayi, Kara, Bachirawa, Yankaba, Mariri and Kabara livestock markets in Municipal, Taruauni, Gmale, Nassarawa, Fagge and Dala Local Government Areas of Kano State. The body weights of sheep were obtained using bathroom scale. Neck length, back length, leg length, height at wither, loin girth and chest girth measurements were taken using a measuring tape. Body weight was measured by measuring the weight of a research assistant while carrying the animal in his hands using bathroom scale. The difference in the human weight from the total weight for each weighing was recorded as the individual animal's weight. The operationalized definitions of categorical observed variables were shown in Table 1.

Table 1: Operationalized definition of categorical observed variables

Variables	Definition
Sex	Male = 1; Female = 2
Breed	Balami = 1; Sudanese = 2; Uda = 3; Yankasa = 4
Body weight class	Small (20-30kg) = 1; Medium (31-40kg) = 2; Large (41-54kg) = 3; Extra-large (55kg and above) = 4

The data obtained were subjected to generalized least squares (GLS) estimator of a triangular seemingly unrelated regression (SUR) model (Lahiri and Schmidt, 1978). Maximum likelihood estimation method of structural equation

modelling was used to generate parameter estimates through recursive system with correlated errors (SEM Command Language) (Stata[®]; Serial number: 501506215411; licensed to FBIW KU Leuven).

Results and discussion

Path analysis is the statistical technique used to examine causal relationships between two or more variables. It is based upon a linear equation system. Path analysis was adopted by the social sciences in the 1960s and has been used with increasing frequency in the ecological literature since the 1970s. In ecological studies, path analysis is used mainly in the attempt to understand comparative strengths of direct and indirect relationships among a set of variables. In this way, path analysis is unique from other linear equation models. In path analysis mediated pathways in those acting through a mediating variable, can be examined. Pathways in path models represent hypotheses of researchers, and can never be statistically tested for directionality. Path analysis is a subset of Structural Equation Modeling (SEM), “allows examination of a set of relationships between one or more independent variables, either continuous or discrete, and one or more dependent variables, either continuous or discrete.”

SEM deals with measured and latent variables. A measured variable is a variable that can be observed directly and is measurable. Measured variables are also known as observed variables, indicators or manifest variables. A latent variable is a variable that cannot be observed directly and must be inferred from measured variables. Latent variables are implied by the covariance among two or more measured variables. They are also known as factors (i.e. factor analysis), constructs or unobserved variables. SEM is a combination of multiple regression and factor analysis. Path analysis deals only with measured variables. The Table 2 below shows the descriptive summary of the data sets used for the study. For breed were used with Yankasa taken with the highest mode, generally male animals were equally the largest category of the animal used, on price the average price of animal was found to be N38,294.13, majority of the animal used had loin girth (loin_gt) of 81.6, chest girth (chest_gt) of 76.42 and height at wither of 72.3.

Table 2: Descriptive summary of data used for the study

Variable	Obs	Mean	Std. Dev.	Min	Max
breed	621	3.243156	1.118959	1	4
sex	621	1.36715	0.482416	1	2
price_2017	621	38294.18	9339.802	16128.04	69106.33
bwt_class	621	2.027375	0.797713	1	4
bwt	621	36.22259	7.815893	22	65.7
neck_lt	621	34.42535	18.06884	20	346
back_lt	621	54.8533	13.47021	22.2	80.8
loin_gt	621	81.58519	12.06101	50	106.6
chest_gt	621	76.4242	10.50836	53	101
height_wht	621	72.31042	8.750533	50.1	92.9
leg_lt	621	44.06031	7.281654	30	62.6

Price_2017 = price of sheep in 2017; bwt_class = body weight class; bwt = body weight; neck_lt = neck length; back_lt = back length; loin_girth = loin_gt = loin girth; chest_gt = chest girth; height_wht = height at wither; leg_lt = leg length

The diagnostic test for the models used in this study were conducted using three major approaches, the model identification, model at a close system and

model at open system. Table 3 shows the standard model identification for the study. Identification involves the evaluation of conditions associated with variables used

Explorative analysis of relationships among breed, growth traits, prices and sex

for modelling in obtaining parameters of each model, which uniquely give solution for every free parameter specified in the model from the observed data. To understand the model three types of model identification presented. If a single, unique value cannot be obtained from the observed data for one or more free parameters, then the model is under identified. Therefore, the solution for y is totally dependent on the solution for x. When there are more unknowns (x and y) than the number of equations (1), the model is under identified. However, if for each free parameter a value can be obtained through one only one manipulation of the observed data, and then the model is just identified. In this, when the number of linearly independent equations is the same as the number of unknown parameters, the model is just identified.

This means we can get unique parameter estimates, but the model cannot be tested. Also, if a value for one or more free parameters can be obtained in multiple ways from the observed data, then the model is over identified. Over identification is the condition in which there are more equations than unknown independent parameters. Based on these conditions from the model below all the parameters of the model were identified, the parameters of their MI values are higher with probability lower than 1% in all cases. This suggested that this model fit well and can be used effectively for estimation purpose. This model is assumed for a closed system. Close system is a system isolated system from its environment. Such a system works under ideal condition where there were no external influences outside the model.

Table 3: The standard model identification

Standard Model Identification					
Measurement	MI	df	P>MI	EPC	EPC
breed Body	61.82	1	0.000	0.40185	2.62172
sex Body	17.601	1	0.000	-0.10038	-1.51921
bwt_class Body	30.35	1	0.000	0.31203	2.85460
bwt type	481.311	1	0.000	47.83290	3.37020
back_lt type	24.008	1	0.000	-17.62315	-0.72123
loin_gt type	39.126	1	0.000	-10.73227	-0.49035
leg_lt type	9.412	1	0.000	-2.97981	-0.22540
cov(e.breed,e.bwt)	6.384	1	0.010	0.36330	0.13334
cov(e.breed,e.back_lt)	62.016	1	0.000	-3.40968	-0.30516
cov(e.breed,e.loin_gt)	24.246	1	0.000	0.97184	0.13913
cov(e.breed,e.chest_gt)	40.437	1	0.000	-0.91983	-0.17951
cov(e.breed,e.height_wht)	12.915	1	0.000	0.42940	0.09217
cov(e.breed,e.leg_lt)	68.949	1	0.000	-0.96039	-0.21445
cov(e.sex,e.back_lt)	8.663	1	0.000	0.61321	0.11350
cov(e.sex,e.height_wht)	4.14	1	0.040	-0.11685	-0.05187
cov(e.bwt_class,e.bwt)	419.866	1	0.000	1.59658	1.38431
cov(e.bwt_class,e.loin_gt)	33.543	1	0.000	-0.49275	-0.16665
cov(e.bwt_class,e.chest_gt)	8.199	1	0.000	-0.18071	-0.08331
cov(e.bwt_class,e.height_wht)	8.511	1	0.000	-0.15030	-0.07621
cov(e.bwt_class,e.leg_lt)	24.169	1	0.000	-0.24163	-0.12745
cov(e.bwt,e.back_lt)	66.928	1	0.000	12.35800	0.38456
cov(e.bwt,e.chest_gt)	18.438	1	0.000	-2.19327	-0.14882
cov(e.bwt,e.height_wht)	16.962	1	0.000	-1.71465	-0.12797
cov(e.bwt,e.leg_lt)	6.105	1	0.010	-0.97879	-0.07599
cov(e.back_lt,e.height_wht)	51.933	1	0.000	-10.46796	-0.19051
cov(e.back_lt,e.leg_lt)	96.033	1	0.000	13.71396	0.25963
cov(e.loin_gt,e.height_wht)	116.862	1	0.000	10.56461	0.30755
cov(e.chest_gt,e.leg_lt)	51.031	1	0.000	5.38460	0.22228

bwt_class = body weight class; bwt = body weight; neck_lt = neck length; back_lt = back length; loin girth = loin_gt = loin girth; chest_gt = chest girth; height_wht = height at wither; leg_lt = leg length

Open system is a system equation, which interact and exchange matters with its environment. Thus, in open system the complex model is a set of interdependent parts, which together make up a whole system in this each parts contribute something and received something from the whole system, which in turn is interdependent with some larger environment. It is a dynamic system, which often is full of surprises and uncertainties. The system keeps evolving through a continuous process of development and strive to attain a state of equilibrium. Table 4 shows the result of the parameters estimation when the system is at its dynamic stage. Imposing constraints at breed type, bwt_body, price and variance of price. All parameters in this model were significant at 1% for actual measured variables, their variance and covariance.

Table 5 shows the parameters of the model at open system. Similarly, in all cases except for bwt_Breed, and neck, all parameters were found to be significant at 1%. The overall saturated model is equally significant at 1% since the probability of the chi square is lower than 1%, which shows the model at saturated stage is in equilibrium and hence all parameters are important in this estimation.

The path diagram for the inter-relationships among breed, sex, prices and linear body measurements of sheep is shown in Figure 1. Price of sheep is an observed variable while type and body parameter are latent variables. Type as a latent variable had direct relationships with breed, sex and body weight class. Body parameter on the other hand had direct relationships with body weight and all the linear body measurements considered in this study. The inter-relationship among price, type and

body parameters are shown to be inter-dependent from one to the other two variables. This indicates that there is feedback from one of the three variables to the others. The assessment showed that the relationship between price and type is negative and high (-0.76). This means that a unit increase in type (combination of breed, sex and body weight class) will bring about lower price of sheep in the market being considered. The relationship between price and body parameters, on the other hand, is positive and high (0.87). An increase in body parameters will bring about an increase in price of sheep. Covariance between body parameters and type was equally positive and high (0.83). Breed and type are positively related (0.63). There is a perfect but negative relationship between body weight class and type (1.0). Sex and type had a positive but low relationship (0.28). Variables with high coefficients among themselves are known to predict one-another close to reality than those with lower coefficients. All the relationships and covariance observed in this study were statistically significant at either 1% or 5% level of significance. Structural equation modeling is used to show how models that had better match the theoretical relations among variables can enhance interpretability and lead to quite different conclusions (Grace and Bollen, 2005).

The z-value for each of the indicator variable is reported in each of the box for that observed indicator thus 2.9 for breed, 2.8 for sex and 2.5 for body weight class, others are 4.6 for birth weight, 1.9 for neck length, 4.1 for back length, 6.8 for loin, 7.4 for chest girth, 8.3 for height at wither and 6.1 for leg length. The seemingly unrelated regression (SUR) models for type, body parameter and price are as shown below:

Explorative analysis of relationships among breed, growth traits, prices and sex

Table 4: Open Information Model

		OIM				
		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
Measurement	breed type	1	(constrained)			
	_cons	3.24396	0.04061	79.89	0.000	3.164376 3.323546
sex	type	0.20219	0.04229	4.78	0.000	0.119312 0.285066
	_cons	1.36731	0.01895	72.14	0.000	1.330162 1.404463
bwt_class	type	-1.36984	0.10955	-12.5	0.000	-1.58455 -1.15513
	_cons	2.02627	0.01843	109.92	0.000	1.990143 2.062403
bwt	Body	1.00000	(constrained)			
	_cons	36.21105	0.15261	237.28	0.000	35.91195 36.51015
neck_lt	Body	0.62066	0.09772	6.35	0.000	0.42913 0.812182
	_cons	34.41818	0.70704	48.68	0.000	33.03241 35.80396
back_lt	Body	0.94813	0.06383	14.85	0.000	0.823031 1.073231
	_cons	54.84236	0.47335	115.86	0.000	53.91462 55.7701
loin_gt	Body	1.23936	0.04577	27.08	0.000	1.149645 1.329072
	_cons	81.57088	0.34478	236.59	0.000	80.89513 82.24663
chest_gt	Body	1.14312	0.03571	32.01	0.000	1.073121 1.21311
	_cons	76.41101	0.27174	281.2	0.000	75.87841 76.9436
height_wht	Body	0.91228	0.03193	28.57	0.000	0.849699 0.974857
	_cons	72.29989	0.23957	301.79	0.000	71.83033 72.76944
leg_lt	Body	0.70832	0.02926	24.2	0.000	0.650967 0.765682
	_cons	44.05213	0.21830	201.8	0.000	43.62427 44.47999
	mean(price)	38278.39000	(constrained)			
	var(e.breed)	0.97598	0.05648			0.871317 1.093208
	var(e.sex)	0.22115	0.01255			0.197865 0.247173
	var(e.bwt_class)	0.12111	0.02190			0.084966 0.172619
	var(e.bwt)	4.88912	0.51819			3.97203 6.017956
	var(e.neck_lt)	306.75350	17.62332			274.0862 343.3144
	var(e.back_lt)	130.53320	7.58010			116.4908 146.2684
	var(e.loin_gt)	59.11414	3.63889			52.3955 66.69431
	var(e.chest_gt)	33.34575	2.14025			29.40407 37.81583
	var(e.height_wht)	27.67541	1.74116			24.46481 31.30735
	var(e.leg_lt)	24.79003	1.50935			22.00146 27.93204
	var(price)	81700000	(constrained)			
	var(type)	0.26010	0.04172			0.189939 0.356167
	var(Body)	53.23287	2.14443			49.19151 57.60625
	cov(e.breed,e.sex)	-0.08993	0.01945	-4.62	0.00000	-0.12804 -0.05182
	cov(e.sex,e.bwt_class)	0.02559	0.00805	3.18	0.00100	0.009812 0.041362
	cov(e.bwt,e.neck_lt)	-5.31019	2.04374	-2.6	0.00900	-9.31583 -1.30454
	cov(e.neck_lt,e.back_lt)	20.54561	8.32202	2.47	0.01400	4.234755 36.85646
	cov(e.back_lt,e.loin_gt)	-11.25335	2.55816	-4.4	0.00000	-16.2673 -6.23945
	cov(e.loin_gt,e.chest_gt)	32.11290	2.57646	12.46	0.00000	27.06314 37.16266
	cov(e.chest_gt,e.height_wht)	3.45838	0.66225	5.22	0.00000	2.160397 4.756371
	cov(e.height_wht,e.leg_lt)	20.49702	1.49722	13.69	0.00000	17.56252 23.43152
	cov(price,type)	-4164.09500	327.29550	-12.72	0.00000	-4805.58 -3522.61
	cov(price,Body)	59733.02000	1330.66400	44.89	0.00000	57124.97 62341.07
	cov(type,Body)	-3.83204	0.32348	-11.85	0.00000	-4.46604 -3.19803

bwt_class = body weight class; bwt = body weight; neck_lt = neck length;
back_lt = back length; loin_girth = loin_gt = loin girth; chest_gt = chest girth;
height_wht = height at wither; leg_lt = leg length

Table 4: Open Information Model (OIM)

Structural	OIM					
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
bwt breed	0.3402899	0.226089	1.51	0.1320	-0.10284	0.783416
sex	-1.37724	0.397168	-3.47	0.0010	-2.15568	-0.59881
neck_lt	0.0055114	0.010289	0.54	0.5920	-0.01465	0.025678
back_lt	0.2479356	0.017202	14.41	0.0000	0.214221	0.281651
loin_gt	0.1355061	0.040858	3.32	0.0010	0.055426	0.215587
chest_gt	0.3488062	0.049442	7.05	0.0000	0.251903	0.44571
height_wht	0.2031603	0.068963	2.95	0.0030	0.067996	0.338325
leg_lt	-0.3623442	0.068149	-5.32	0.0000	-0.49591	-0.22877
_cons	-13.22609	2.747329	-4.81	0.0000	-18.6108	-7.84143
var(e.bwt)	19.59385	1.11196			17.53129	21.89908

LR test of model vs. saturated: $\chi^2(0.000) = 0.00$, Prob > χ^2

bwt = body weight; neck_lt = neck length; back_lt = back length; loin_gt = loin girth; loin_gt = loin girth;
chest_gt = chest girth; height_wht = height at wither; leg_lt = leg length

Figure 1: Path diagram of direct and indirect relationships among observed and latent variables

$$\begin{aligned}
 type &= 1 + 0.63breed_1 + 0.28sex_2 + 1.0bwt_classX_3 \\
 body_param &= 1 + 0.9bwt + 0.28neck_lt_1 + 0.6back_lt + 0.85loin_gt + 0.91chest_lt \\
 &\quad + 0.88heightwt + 0.9leglt \quad price = 1 - 0.76type_1 + 0.83body_param_2
 \end{aligned}$$

The variance and covariance parameters values and significant are very important indicators in determining how successful this price estimation study can be conducted. Variance in statistics is a measurement of the spread between numbers in a data set. That is, it measures how far each number in the set is from the mean and therefore from every other number in the set. Variance, covariance and correlation are the basic parameters in this price of sheep estimation analysis.

The covariance shows the degree of the joint variation of two random variables. More values of one variable predominantly resemble with the greater values of the other variable, if two variables move in the same direction either increasing or decreasing they are said to positive or negatively covary whereas if one increase the other decrease then they are said to be inversely related. In the opposite case, when the greater values of one variable mainly correspond to the lesser values of the other, (i.e., the variables tend to show opposite behavior), the covariance is negative. The sign of the covariance therefore shows the tendency in the linear relationship between the variables. The magnitude of the covariance is not easy to interpret because it is not normalized and hence depends on the magnitudes of the variables. The normalized version of the covariance, the correlation coefficient, however, shows by its magnitude the strength of the linear relation. In sheep price estimation the covariance might have either positive or negative tendency if two parameters are linked the tendency to estimate accurately the price sheep is possible whereas if they are inversely related, in this case use of such parameters may likely provide a problematic pricing process.

Conclusion

Simultaneous prediction equations for

estimating prices of sheep had been generated for some selected livestock markets in Kano State.

References

- Grace, J. B. 2006.** Structural equation modeling and natural systems. Cambridge University Press, Cambridge, UK.
- Grace, J. B. and Bollen, K. A. 2005.** Interpreting the results from multiple regression and structural equation models. *Bulletin of the Ecological Society of America*, 15(5), 283-295.
- Iyiola-Tunji, A. O., Olugbemi, T. S., Ali, A. O. and Ojo, O. A. 2011.** Inter-relationship between body measurements and prices of sheep in an open market in Kano State. *Animal Production*, 13(1), 64-68.
- Lahiri, K. and Schmidt, P. 1978.** A Note on the Consistency of the GLS Estimator in Triangular Structural Systems. *Econometrica*, 46: 1217-1221.
- Schumaker, R. E. and Lomax, R. G. 2010.** *A beginner's guide to structural equation modelling*. Taylor and Francis Group, LLC. Third Edition. New York. 510pp. ISBN: 978-1-84169-890-8.
- Yakubu, A. 2011.** Path analysis of conformation traits and milk yield of Bunaji cows in smallholder's herds in Nigeria. *Agricultura Tropica et Subtropica*, 44(3): 152-157.

Received: 31st August, 2019

Accepted: 20th February, 2020