HOUSEHOLD MILK CONSMPTION IN POLAND: AN APPLICATION OF BIVARIATE TWO-PART MODEL

Shengfei Fu[1], Anna M. Klepacka[2], and Wojciech J. Florkowski[3]

[1] PhD student, University of Georgia, Department of Ag & Applied Economics.
306 Conner Hall, Athens GA 30602   Email:shengfei@uga.edu

[2] Faculty of Production Engineering, Warsaw University of Life Science-SGGW
164 Nowoursynowska St., 02-787 Warsaw, Poland   Email:annaklepacka@hotmail.com

[3] Professor, University of Georgia, Department of Ag & Applied Economics.
1109 Experiment St., 212 Stuckey Building, Griffin, GA 30223   Email:wojciech@uga.edu

Poster paper prepared for presentation at the EAAE 2014 Congress
'Agri-Food and Rural Innovations for Healthier Societies'

August 26 to 29, 2014
Ljubljana, Slovenia

HOUSEHOLD CONSMPTION IN POLAND: AN APPLICATION OF BIVARIATE TWO-PART MODEL

Shengfei Fu, Anna M. Klepacka, and Wojciech J. Florkowski

Abstract: Milk products are a preeminent food category in Poland, providing both employment and dietary benefit. This paper investigates factors affecting household milk consumption in Poland, paying attention to the effect of outmigration. For the analysis of consumption decisions on whole and low fat milk, bivariate two-part model is selected after a comparison to multivariate sample selection model.

Keywords: milk consumption, outmigration, bivariate two-part model, multivariate Heckman sample selection model.

Introduction

In Poland, the dairy sector is one of the most important parts of the food industry (Sznajder, 2012). Given milk's nutritional values and the importance of diary production, processing, retailing and consumption in Poland, this paper investigates demographic, socio-economic and location factors, and their connection to milk product consumption. Special attentions are paid to the effect of worker outmigration, a current issue in Poland.

Because of the specificity of dairy production, expanding milk production in regions with suitable natural conditions could provide job opportunities in rural areas (Klepacka et al., 2013). An analysis of factors influencing dairy consumption, including milk of various fat contents, also offers insights applicable in dairy processing and distribution, and, even in assessment of potential public health threats resulting from permanent changes in dairy product consumption.

Methods

Two censored systems, multivariate Heckman's sample selection model (MSSM) (Steward and Yen, 2004; Yen, 2005) and its nested form bivariate two-part model (B2P) are considered. Such system approaches allow the correlation between the consumption decisions of whole and low fat milk. Each outcome variable y_i (milk expenditure) is governed by a binary selection rule of whether to consume as follows, where observation subscription is omitted for convenience:

$$\log(y_i) = x'\beta_i + v_i \quad \text{if } z'\alpha_i + u_i > 0 \tag{1}$$

$$y_i = 0 \quad \text{if } z'\alpha_i + u_i \leq 0 \quad i = 1, 2$$

where z and are x vectors affecting binary purchase decision and quantity decision, respectively; α and β are vectors of parameters; $\boldsymbol{u_i}$ and $\boldsymbol{v_i}$ are random error in the participation and level equation, respectively. The concatenated error terms are assumed to follow multivariate normal distribution with mean zero and covariance maxtrix invlolvng the standad deviation of $\boldsymbol{v,}$ and the correlation matrices among elements of $\boldsymbol{u}$ and $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{u}$, and $\boldsymbol{v}$ and $\boldsymbol{v}$, respectively.

The MSSM assumes the concatenated error vector distribted as 4-variate normal. The B2P model assumes no correlation between the participation and level equation, therefore it essentially constitutes a bivariate probit regression and a bivariate lognormal regression. Maximum likelihood estimation (MLE) are applied to obtain parameter estiamtes.

Model Selection

There is a well-established debate in health econometrics over the merits of Heckman sample selection models versus two-part models. We choose the B2P model after a comparison based on theoretical and statistical grounds (Madden, 2008). Firstly, the two-part model is more appropriate for the purpose of analyzing the actual spending instead of the potential expenditure. Additionally, we estimate both MSSM and B2P model and compare each expanatory variable's empirical mean square error (M.S.E.) (Leung and Yu, 1966; Madden, 2008). For each variable, the model gives lower M.S.E. is selected.

Marginal Effects

For probit regression, the signs of parameter estimates reveal the direction, but not the magnitudes of their effects. A more economically meaningful measure, marginal effects, are calculated based on conditional means for the joint distribution. Elasticity is defined as percent change in probability of purchase, corresponding to a one-unit change in explantory variable.

The marginal effects for indicator variables are obtained by differencing purchase probability when the explanatory variable takes value of one versus zero. For the lognormal regression, the parameter estimate itself is the marginal effect.

Data, Sample and Variables

The data are from the Polish household panel of about 20,000 households annually surveyed by Poland's National Statistics Office (GUS). We use a pooled cross-sectional sample of 108,064 non-missing observations for the period of 2004 to 2008.

Table 1. Variable definitions and sample statistics (n=108,064)

Variable	Description/Unit	Mean	SD
Food Expenditures / Dependent variables			
Buy1	1, if household buys whole milk, 0 otherwise	0.612	0.487
Buy2	1, if household buys low fat milk, 0 otherwise	0.624	0.484
Wmilk	Expenditure on whole milk in the month preceding survey, in Polish Zloty (PLN)	19.644	22.204
Lmilk	Expenditure on low fat milk in the month preceding survey (PLN)	14.439	14.700
Demographic, Socio-Economic Factors/Explanatory variables			
Village	1, if a household residents in village, 0 otherwise	0.362	0.481
Income	Household income in the month preceding survey (PLN)	2.306	1.473
Male	1, if the household head is male, 0 otherwise	0.595	0.491
Married	1, if the household head is married, 0 otherwise	0.678	0.467
Educ	1, if the household head has secondary or higher education, 0 otherwise	0.402	0.490
Age	Household head's age, in years	50.888	15.288
Employ	1 if household head is permanently employed or contract employee, 0 otherwise	0.421	0.494
Child	Number of children (under 18)	0.639	0.985
Adult	Number of adults 60 or under 60 years old	1.823	1.198
Elder	Number of elders above 60	0.451	0.695
OutD	Net migration domestically to other regions in Poland, in 1000	-1.252	5.598
OutF	Net migration international to other countries, in 1000	1.333	2.053
Time	1-5 corresponding to year 2004 - 2008	2.762	1.473

Table 1 presents variable definitions and sample statistics. Average milk expenditures and usual demographic factors are reported. In addition, two variables are reported as measure of depopulation. First, net domestic migration measures the net outflow of population from a region to other regions within Poland. Second, net international migration measures the net outflow of population from a region to other countries

Results and Discussions

Both MSSM and B2P model are estimated. Marginal effects on the probability of purchase and on the conditional mean expenditure are calculated.

Table 2. Marginal effects, bivariate two-part model

Variable	Elasticity of Purchase Probability		Elasticity of Conditional Mean Expenditure	
	Whole milk	Low fat milk	Whole milk	Low fat milk
	Est. (Asy.se)	Est. (Asy.se)	Est. (se)	Est. (se)
Constant	-0.149***	0.190***	1.712***	1.553***
	(0.018)	(0.011)	(0.024)	(0.023)
Income	0.005	0.022**	-0.012***	0.024***
	(0.008)	(0.010)	(0.003)	(0.003)
Educ	-0.010***	-0.004	-0.034***	0.019**
	(0.003)	(0.004)	(0.009)	(0.008)
Married	0.034***	0.042***	0.048***	0.099***
	(0.005)	(0.005)	(0.010)	(0.010)
Child	0.052***	0.005	0.197***	0.172***
	(0.003)	(0.004)	(0.004)	(0.005)
Adult	0.052***	-0.011	0.17***	0.105***
	(0.005)	(0.009)	(0.004)	(0.005)
Elder	0.081***	-0.029***	0.268***	0.176***
	(0.006)	(0.006)	(0.008)	(0.009)
Age	-0.001	0.003	-0.001***	0.004***
	(0.173)	(0.213)	(0.00)	(0.00)
Male	-0.001	-0.032***	0.067***	0.021***
	(0.004)	(0.005)	(0.009)	(0.009)
Employ	-0.018***	0.042***	-0.145***	-0.040***
	(0.003)	(0.004)	(0.009)	(0.009)
Village	0.123***	-0.171***	0.486***	0.091***
	(0.003)	(0.004)	(0.009)	(0.009)
OutD	-0.012***	0.010**	-0.001*	0.011***
	(0.004)	(0.006)	(0.001)	(0.001)
OutF	0.018***	-0.008*	-0.019***	-0.036***
	(0.005)	(0.006)	(0.002)	(0.002)
Time	0.025***	-0.027**	0.021***	0.001
	(0.009)	(0.012)	(0.003)	(0.003)
	-0.578***		-0.128***	
Rho	(0.006)		(0.012)	

Note: *** significant at 1%; ** significant at 5%; * significant at 10%.
The asympototic standard errors are obtaine by the delta method (Spanos, 1999).

In order to compare the two models, we estimate the empirical M.S.E. under two different null hypotheses: first on the basis that the true model is the MSSM and second, that the true model is the B2P model. Both tests conclude the B2P model is the "true" model.

The botton line of Table 2 presents the estimated correlation coefficients for the B2P model. Consistent to expectation, the decision to buy the two milk products under study and the expenditure amounts are negatively correlated, as reflected by the negative correlation coefficient estimates of -0.578 and -0.128, respectively.

As shown in Table 2, the signs of all variables are not the same across the participation equation and the level equation, confirming that different rules are applied in the two decision stages.

In the binary decision to buy milk, income positively influences the probability of buying both whole and low fat milk. For example, the probability of buying low fat milk grows by 2.2% when income increases by1000PLN. *Ceteris paribus*, more affluent households spend more on low fat milk, but less on whole milk. Specifically, if household income increases by 1000PLN, the expenditure on whole (low fat) milk declines (increases) by 1.2% (2.4%).

Household heads with higher education are more likely to buy low fat milk. They also spend more on this product, compared to their counterpart with lower education attainment level. Low fat milk has the same nutritional benefits as whole milk but contains less fat and is generally considered healthier. People with higher education are more likely to recognize the relative superiority of low fat milk and, therefore, spend more on it.

Married household heads are more likely to buy, and spend more as well, on both milk products. Older household heads are associated with higher expenditure on both milk products, as older population generally consumes more dairy products.

The effect of more family members (children, adults, or elders, respectively) is estimated to increase the probability of buying whole milk. Conditional on purchase, more family members are related with higher expenditure on both milk products.

Male household head are less likely to buy low fat milk. This is consistent with findings reported in literature that females are generally more concerned about diet healthiness, especially with regard to fat intake (Wardle et al. 2004). However, once they have decided to buy milk, male household heads on average make larger purchase.

Households with relatively stable employment are less likely to buy whole milk, but more likely to buy low fat milk. And employment stability is related to lower milk expenditures, possibly because these household heads make fewer shopping trips, or because these households eat away-from-home more often.

Households residing in villages have higher probability of buying whole milk and lower probability of buying low fat milk, reflecting relatively inferior milk choice in rural area. Conditional on purchase, however, rural residents spend more on both milk products, compared to their urban counterparts.

Worker outmigration shows an interesting effect. Houses residing in regions with higher domestic outmigration are less likely to buy whole milk and also spend less on it. As for low fat milk, higher domestic outmigration relates to higher purchase likelihood and higher expenditure. In contrast, higher international outmigration is associated with higher (lower)

probability of buying whole (low fat) milk. Conditional on purchase, households in regions with higher international outmigration spend less on both milk products.

Since regions with high outmigration tend to be economically less developed, the supported development of diary sector (at least milk products) in these areas may provide the potential of additional employment and economic development.

In summary, our model estimation enables a close scrutiny on the demographic and socio-economic factors affecting household whole and low fat milk consumption in Poland. The resulted findings revealed the direction of each variable's effect as well as its magnitude. The findings are important to learn about which factors are associated with healthy (unhealthy) milk choice and are informative for the formulation of economic and public health policies. The estimated effect of depopulation implies the importance of fluid milk in diets of regions loosing residents. Also, the study provides a notion about potential of diary sector (milk products) for local employment and economic development.

References

Klepacka A. M., T. S. Sobczynski, W. J. Florkowski, and C. Revoredo-Giha. (2013). Effects of the Common Agricultural Policy a Non-Family Farm Employment in Primary and Secondary Agricultural Areas. Paper presented at the joint AAEA/CAEA annual meeting, Washington, D. C., August 4-6, 2013.

Leung, S.F. and S. Yu. (1996). On the choice between sample selection and two-part models. *Journal of Econometrics* 72: 197–229.

Madden, D. (2008). Sample selection versus two-part models revisited: the case of female smoking and drinking. *Journal of Health Economics* 27: 300-307.

Spanos, A. (1999). *Probability Theory and Statistical Inference: Econometric Modeling with Observational Data*. Cambridge, UK: Cambridge University Press.

Stewart, H. and S. T. Yen. (2004). Changing household characteristics and the away-from-home food market: a censored equation system approach. *Food Policy* 29(6): 643–658.

Sznajder, P. (2012). Determinanty polskiego eksportu produktów mleczarskich. *Roczniki Naukowe SERiA*, XIV (1):514-518.

Wardle, J., Haase A.M., Steptoe A., Nillapun M., Kiriboon K., and Bellisie F. (2004). Gender differences in food choice: The contribution of health beliefs and dieting. *Annals of Behavioral Medicine* 27(2):107-116.

Yen, S. (2005). A Multivariate Sample-selection Model: Estimating Cigarette and Alcohol Demands with Zero Observations. *American Journal of Agricultural Economics* 87(2): 453-466.