CONSUMPTION OF MUSHROOMS: A DOUBLE-HURDLE MODEL APPROACH

Yuan Jiang, Graduate Student, Department of Food and Resource Economics, University of Florida, E-mail: jyspring@ufl.edu

Lisa House, Professor, Department of Food and Resource Economics, University of Florida, E-mail: lahouse@ufl.edu

Christian Tejera, Graduate Student, Department of Food Science and Human Nutrition, University of Florida,Email: christej@ufl.edu

Susan S. Percival, Professor, Department of Food Science and Human Nutrition, University of Florida, Email: percival@ufl.edu

Selected Poster prepared for presentation at the Southern Agricultural Economics Association's 2015 Annual Meeting, Atlanta, Georgia, January 31-February 3, 2015

Abstract：

This study investigates demographic, socioeconomic and food attributes factors contributing to the consumption of fresh and processed (jarred/canned) mushrooms. The double-hurdle model is employed to analyze the data. Awareness of mushroom health benefits, mushroom's attributes, income, household size, race and age are significant determinants of the mushroom consumption.

Key Word: Fresh and processed mushroom consumption, double-hurdle model

Consumption of Mushrooms: A double-hurdle Approach

Introduction

The United States is the world's second-largest producer of mushrooms, with 16 percent of world output, following China, and in terms of value of production, mushrooms are a leading U.S specialty crop, exceeded only potatoes, tomatoes and lettuce. On the side of consumption, in the past decades, more and more health-conscious consumers increasingly realize the importance of consuming mushrooms. Consumption of mushrooms has been on the rise in the United States over the past several decades. Typically used as a vegetable, per capita consumption of this carefully cultivated fungus crop has quadrupled since 1965 (the first year for which reliable data are available). According to data compiled by the U.S Department of Agriculture's (USDA) Economics Research Service (ERS), per capita use of all mushrooms totaled about 4.0 pounds in 2011, compared with about 0.69 pounds in 1965. (Figure 1)

Fresh-market mushroom account for three-fourths of domestic consumption in the year of 2011. Since the year of 1990, per capita consumption of fresh mushrooms has increased dramatically, while interestingly, per capita consumption of processing mushrooms (mostly canned mushrooms) has gradually declined since then. The consumption of fresh mushrooms is only a litter greater than that of canned mushrooms at the year of 1990, it becomes three times higher than the consumption of canned mushrooms since the year of 2008.

Given this deviation in the consumption of the two types of mushrooms, it is interesting to investigate the potential differences existing in the two markets. Thus it may provide the industry a way to understand the consumption patterns of different types of mushrooms, thus developing different production strategies and marketing strategies.

There has been a lot of researches investigating both the nutritional and medical benefits of mushrooms. In recent years there has also been continual interest in information regarding the consumption distribution of foods such as mushrooms. A study completed by the University of Alabama (Onianwa et al., 2000) analyzed the market potential for locally grown shiitake mushrooms at the retail-level in northern Alabama. The North Carolina Department of Agriculture and Consumer Services

performed a survey (Augostini, 2002) to identify the trends in mushroom demand for the next 5 years. The target audience for this survey included retail and wholesale produce buyers from North Carolina, South Carolina, and Virginia. It demonstrated that produce buyers are optimistic that demand for specialty mushrooms will grow over the next five years. In 2006, the University of Missouri Center for Agroforestry conducted a nationwide survey of shiitake mushroom producers to analyze the U.S. shiitake mushroom industry, and found out both the barriers and opportunities for famers in the mushroom market (Gold et al.,2008).

Although much is known about the supply side of the U.S. mushroom market, relatively little detailed information has been published concerning consumer demand, aside from such things as basic USDA disappearance data and retail sales information. There is few research exploring factors influencing mushroom consumption at individual level. The research conducted by Lucier in 2003 demonstrated that Asian and non-Hispanic white consumers were the strongest consumers of mushrooms; per capita mushroom consumption was positively correlated with income; men and women between 20 and 39 years old were the leading mushroom consumers, representing about 32% of the population, yet consuming 43% of all mushrooms (Lucier et al.2003).

The purpose of this research is to investigate the factors that affect both fresh and processing mushroom consumption through a survey conducted in 2012.The analysis attempts to explain seafood consumption by consumers' characteristics such as their demographic and socioeconomic conditions. By analyzing the results, we further compare the differences between in fresh and processing mushroom market. At last policy implications on mushroom producers, retailers, importers, and policy makers are drawn based on the findings in this research.

Model

Since the consumption of fruits demonstrates substantial diversity and seasonality, it's common to have many zero observations during certain periods. The Ordinary Least Square method will no longer the proper method to get the consistent by excluding these zero observations (William H. Greene, 2007). Basically, there are three reasons caused

the zero observation (Humphreys,2013): non-consumption, corner solution or purchase infrequency. Non-consumption happens when people choose not to purchase because of some stable reasons like preference and diet habits. Corner solution specify that people choose not to consume because of expenditure budget, income or prices. Infrequency was mostly happen for durable goods, that no purchase behavior were observed during a certain period of time.

Dealing with the problem of "many zeros", there have been many approaches and one of the most widely used approach is the Tobit model (Tobin, 1958). It was developed to alleviate the problems caused by OLS. However, it is still very restrictive by assuming variables which determine the probability of consumption also determine the frequency of consumption. To solve the shortcoming of the Tobit model, a number of generalizations to the Tobit model has been developed. The double-hurdle model was one of the generalizations which was used widely to solve the problem of "excess zeros". The Cragg's independent model (Cragg, 1971), which is a double-hurdle model, relaxes the Tobit model by allowing separate stochastic processes for the participation and consumption decisions (Yen and Huang, 1996). In Double hurdle, the purchase behavior was divided into two stages: the participation stage and consumption stage. In the first stage, the consumers make decisions on whether to purchase or not. In the second stage, a decision on how much/many to purchase was determined. Using the double-hurdle will allow us to discuss the two different stages of purchasing

The model specification for the hurdle models are as follows:

The participation stage was defined as:

$$D_i = Z_i\alpha + \mu_i \qquad D_i = \begin{cases} 1 \; if \; Z_i\alpha + \mu_i > 0 \\ 0 \; if \; otherwise \end{cases} \qquad (1)$$

And the consumption stage is:

$$Y_i^* = X_i\beta + \varepsilon_i > 0 \qquad (2)$$

$$Y_i = 0 \quad \text{if } \{Z_i\alpha + \mu_i \le 0 \,\&\, X_i\beta + \varepsilon_i > 0\} \qquad (3)$$

$$\text{or } \{Z_i\alpha + \mu_i > 0 \,\&\, X_i\beta + \varepsilon_i \le 0\} \qquad (4)$$

$$\text{or } \{Z_i\alpha + \mu_i \le 0 \,\&\, X_i\beta + \varepsilon_i \le 0\} \qquad (5)$$

$$Y_i = Y_i^* > 0 \text{ if } \{Z_i\alpha + \mu_i > 0 \,\&\, X_i\beta + \varepsilon_i > 0\} \quad (6)$$

Where Di is the latend binary dependent variable indicating whether or not participation occurs. Yi is the observed purchase frequency and Yi* is the latent consumption variable. Zi and Xi are the vectors of explanatory variables in the two hurdles respectively. α and β are the vectors of coefficients to be estimated. μi and εi are the error terms. The μ_i and ε_i have the distribution:

$$\begin{pmatrix}\mu_i \\ \varepsilon_i\end{pmatrix} = \mathrm{N}\left(0, \begin{pmatrix}1 & \rho\sigma \\ \rho\sigma & \sigma^2\end{pmatrix}\right) \quad (7)$$

Where ρ is the correlation coefficient between μ_i and ε_i.

As described above, zero was allowed in both of the two stages, and a positive consumption frequency was observed only when the two hurdles are crossed.

The likelihood function under dependence can be expressed as:

$$\prod 0\,[1 - \Pr(Y_i^* > 0, D_i > 0)] * \prod + \Pr(Y_i^* > 0, D_i > 0)\Pr(Y_i^* | Y_i^* > 0, D_i > 0) \quad (8)$$

With the dependent variable being purchase frequency, a Poisson distribution is appropriate. While as described above, the correlation between the two equations ρ could possibly exists, since both of the two purchase stages are determined by the same person. However, whether it was truly exist or not, the further statistic test will be employed. The parameter for the Probit of the first hurdle is denoted as θ and that for the Poisson of the second hurdle is denoted as λ. Based on Holgate's bivariate Poisson distribution, the log-likelihood function is (Shonkwiler and Shaw, 1996)

$$LL = \sum_0 \ln[\exp(-\lambda - \rho) + \exp(-\theta - \rho) - \exp(-\lambda - \theta - \rho)] +$$

$$\sum_+ \ln[\frac{\exp(-\lambda - \rho)(\lambda + \rho)^Y}{Y!} - \frac{\exp(-\lambda - \theta - \rho)\lambda^Y}{Y!}]$$

Where ρ is the covariance parameter, θ=exp($Z_i\alpha$) and λ= exp ($X_i\beta$), α and β can be estimated using maximum likelihood estimation in both models.

Independent variables in our study are diet constraint, household size, sex, age, racial group, level of education, employment status, household's annual income, awareness of the benefits of mushroom, characteristics of mushrooms; while the

dependent variables is whether they purchase mushroom, and purchase frequency in the last month.

Data

An online survey about consumers' mushroom consumption were conducted with random panel of respondents. It was launched on September 17, 2012 at 5:34 PM and closed on September 20, 2012 at 10:48 AM. A total of 1217 responds were collected. The target respondents are consumers in the northeastern and southeastern states of the United States. Respondents answer a series of questions on how often and why (or why not) they purchase different types of mushrooms.

Due to the difficulty in collecting the price information across different purchase locations, purchase frequency information is used to represent the consumption amount for each household. In the survey, we first asked whether the respondent had ever purchased fresh or canned/jarred mushroom and then asked whether they had purchased fresh or Canned/Jarred mushroom in the month before the survey was taken. For those respondents who had purchased in the prior month, we asked them how many times they purchased, separately for fresh and canned mushroom. We only asked purchase information for the month before the survey was taken to ensure accuracy of the data as it is usually difficult for people to recall purchases more than one month ago.

Demographic information was also collected. For the participation stage, we include diet constraint, income, race, and gender, and education level, age, household size, awareness of the health benefits of mushrooms, characteristics of mushrooms as the explanatory variables. For the consumption stage, the same set of explanatory variables were used in the participation stage. (Table 1&Table 2).

Results and Discussion

Among the survey, 54.9% of the participants are female, it might be a result by setting the primary grocery shopper criteria. Caucasians substitutes the main majority of the participants (75%), followed by African-American/Black (11.1%), Asia (7.7%), Hispanic (7.5%). 77.0% of the participants have some college or a four-year degree.15%

have post-graduate degrees. The average age of the participants is 45 years old. The average household income is approximate $42500 per year.

According to the dataset, 52.60% of the participants reported 0 canned mushroom purchase, and 36.99% of the participants reported 0 purchase of fresh mushroom. This result is consistent with the data from USDA, that the consumption of processed mushroom has been lower than the consumption of fresh mushrooms. The estimated result for fresh mushroom were displayed in Table 3. For Fresh mushrooms, income, race, age, sex, characteristics of mushrooms, food expenditure budget, and awareness of health information are all significant in the first stage. Males are more likely to purchase fresh mushrooms than females. People with higher income or with higher food expenditure budget would be more likely to buy fresh mushrooms. Caucasian and Black are less likely to decide to buy the fresh mushrooms. People between 18 to 29 and people fall in the age range of 30-49 would like to more likely to buy fresh mushrooms. More important, the result indicated that people who aware of the benefit of mushrooms are more likely to decide to purchase the fresh mushrooms. As expected, fresh mushroom in better condition would attract consumers to buy them, and the most important characteristics are freshness, color and diversity of types available. In the second stage, people who is on diet or trying to lose weight would like to purchase more than other people. Caucasian and Black would buy less than fresh mushrooms. People between 18 to 29 years old would also purchase more. Again, food expenditure budget is significantly positive related with the frequency of buying fresh mushrooms. The awareness of mushroom benefits would again stimulates to buy more fresh mushrooms. Households of 5 people or more is also significantly positive related to the frequency of buying fresh mushrooms. Without any doubt, the better conditions of the mushrooms would attract people to buy more.

For canned/jarred mushrooms, in the first stage (decide to purchase or not), budget is again significantly positive, however, income is no longer significant. The result makes sense because the canned/jarred mushrooms are always less expensive than fresh mushroom. People between 18-29 years old were likely not to buy the canned/jarred mushroom. It also makes sense because jarred/canned mushrooms might not be considered as healthy as fresh mushrooms. One more interesting result is that although

income is not significant, the sign is negative here, which might be an indicator that people with higher income might not purchase processed mushrooms. Hispanic was more likely to buy canned/jarred mushrooms, and people with awareness of mushroom health benefits would be more likely to buy. This time, the characteristics of mushroom are no longer significant as expected. Since, for canned/jarred mushrooms, it would be more difficult to observe the characteristics. In the second stage, Hispanic, awareness of health benefits, budget are significantly positive related with the purchase frequency of canned/jarred mushrooms, while people fall in the range of 18-29 years old would likely to buy less than other people.

The results showed that the awareness of health benefits of mushrooms plays a very important role influencing people's purchase behavior. Thus, it is a wise method to increase mushroom volume by providing the knowledge of mushroom health benefit to more consumers. It indicated that youngsters prefer to buy more fresh mushrooms, while refused the purchase of processed mushrooms. The characters of fresh mushrooms are also important during consumers' purchase process. To always keep mushrooms in good conditions will attract people to buy more.

Besides, covariance between two stages were significant for both two types of mushrooms, which are consistent with our expectation. Both of the two decision stages were made by the same participants, which explained the existence of the correlation between the two equations.

Conclusion:

Since non-consumption of certain food items is very common for individual level survey data, the specification of double-hurdle model provide an appropriate way to deal with these "excess zeros". The results from this study also indicated that the factors influencing participation stage (decide whether to purchase or not) are different from the factors influencing consumption stage (decide how much/many to purchase). Employing the double-hurdle model would provide us more information, which will be valuable to analyze and understand the consumers' purchase behavior, thus provide valuable information for retailers, producers and policy makers to develop appropriate marketing strategies and programs.

Results presented above suggests that the attributes of fresh mushrooms and people's awareness of mushroom benefits dominated consumer's decisions in both of the two stages. The producers and retailers would better target their producing and marketing strategies in these two areas.

References

Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica: Journal of the Econometric Society*, 829-844.

Gold M A, Cernusca M M, Godsey L D. A competitive market analysis of the United States shiitake mushroom marketplace[J]. *HortTechnology*, 2008, 18(3): 489-499.

Jones, A. M. (1989). A double-hurdle model of cigarette consumption. *Journal of applied econometrics*, 4(1), 23-39.

Humphreys, B. R. (2010). Dealing with zeros in economic data.

Lucier, G., Allshouse, J. E., & Lin, B. H. (2003). Factors affecting US mushroom consumption. US Department of Agriculture, *Economic Research Service*

Lin, C. T. J., & Milon, J. W. (1993). Attribute and safety perceptions in a double-hurdle model of shellfish consumption. *American Journal of Agricultural Economics*, 75(3), 724-729.

McDowell, A. (2003). From the help desk: hurdle models. *The Stata Journal*,3(2), 178-184.

Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of econometrics*, 33(3), 341-365.

Ridout, M., Demétrio, C. G., & Hinde, J. (1998, December). Models for count data with many zeros. In Proceedings of the *XIXth International Biometric Conference* (Vol. 19, pp. 179-192).

Shonkwiler, J. S., & Shaw, W. D. (1996). Hurdle count-data models in recreation demand analysis. *Journal of Agricultural & Resource Economics*, 21(2).

Shi, L., House, L., & Gao, Z. (2011, May). Consumer Structure of the Blueberry Market: A Double Hurdle Model Approach. *In 2011 Annual Meeting, July 24-26, 2011, Pittsburgh, Pennsylvania (No. 103630).* Agricultural and Applied Economics Association.

Yen, S. T., & Jones, A. M. (1997). Household consumption of cheese: an inverse hyperbolic sine double-hurdle model with dependent errors. *American journal of agricultural economics*, 79(1), 246-251.

Figure 1. Mushroom Consumption Changes

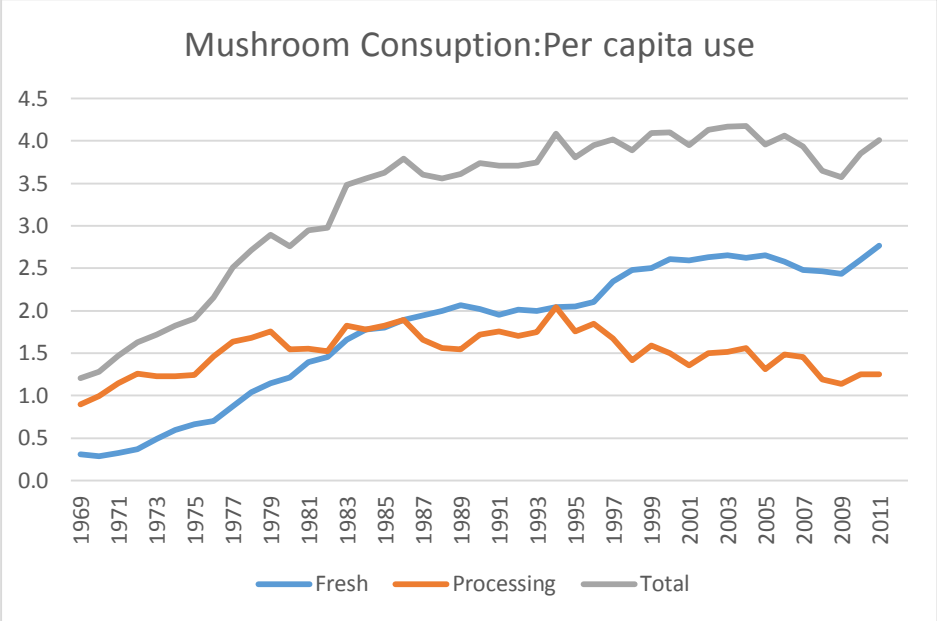

Table 1. Statistics of Demographic Information

Variable	Variable Description	Sample%(N=703)
Gender	Female	54.91
	Male	45.09
Age	<=30	6.75
	31-49	70.55
	>49	22.7
Education	Less than Some College	23.04
	Some College or Bachelor	61.88
	Postgraduate	15.08
Number of Kids	1	15.34
	2	53.37
	≥3	31.29
Race	White	11.10
	Black	75.53
	Hispanic	7.40
	Asia	7.68
	Other race	2.7
Employment	Full-employment	36.98
	Part-time	17.50
	Other	45.52
Diet	Vegan	6.54
	Loss Weight	4.23
	Restrict	23.04
	Normal	66.19
Awareness of mushroom benefit	Yes	59.6
	No	40.40
Family Size	1	56.76
	2	33.85
	≥ 3	9.39
# of kids	1	65.72
	2	17.07
	≥ 3	17.21

Table 2. List of Explanatory Dummy Variables and Explanation

Explanatory Variable	Explanation
diet_vegan	=1 if vegetarian or vegan; =0 otherwise
diet _wloss	=1 if on diet for weight loss;=0 otherwise
caucasian	=1 if Caucasian; =0 otherwise
black	=1 if Black or African American; =0 otherwise
hispanic	=1 if Hispanic; =0 otherwise
asian	=1 if Asian; =0 otherwise
male	=1 if male; =0 if female
college	=1 if education level is some college or higher; =0 otherwise
awarehealth	=1 if aware of the health benefits of mushrooms; =0 otherwise
age18_29	=1 if of age between 18 and 29; =0 otherwise
age30_49	=1 if of age between 30 and 44; =0 otherwise
age45_59	=1 if of age between 45 and 59; =0 otherwise
people3_4	=1 if household has 3 or 4 people; =0 otherwise
people_5above	=1 if household has 5 people or above; =0 otherwise

Table 3. Estimation Results of Double-Hurdle Model

	Explanatory Variable	Fresh Mushroom		Canned Mushroom	
		Coefficient	P-value	Coefficient	P-value
Participant Stage	diet_vegan	.056	.273	.136	.223
	diet _wloss	.226	.165	.040	.148
	income	**.122**	**.029**	**-.051**	**.027**
	caucasian	**-.570**	**.215**	.253	.170
	black	-.506	.282	.228	.236
	hispanic	.195	.326	.488	.252
	male	**.237**	**.127**	-.0619	.117
	college	-.083	.156	-.075	.142
	Awarehealth	**.302**	**.129**	.161	.118
	age18_29	**.412**	**.191**	**-.446**	**.178**
	age30_49	**.308**	**.154**	-.169	.143
	peop3_4	.154	.164	.217	.150
	peop_5above	.487	.351	.258	.271
	budget	**.129**	**.063**	**.158**	**.056**
	color	**.142**	**.054**	.018	.048
	size	.012	.068	.036	.063
	fresh	**.261**	**.075**	-.373	.174
	price	**-.209**	**.065**	.074	.056
	availability	**.134**	**.066**	.043	.061
	constant	**-1.091**	**.475**	.720	.468
Consumption Stage	diet_vegan	.123	.200	.191	.271
	diet _wloss	**.404**	**.128**	.151	.184
	income	-.005	.024	-.052	.035
	race_caucasian	**-.433**	**.150**	.181	.211
	race_black	**-.375**	**.209**	.064	.291
	race_hispanic	.206	.216	**.626**	**.296**

race_asian	0.089	0.027	0.189	0.003
male	.130	.104	-.048	.146
educ_college	-.010	.126	-.203	.1762
Awarehealth_yes	**.314**	**.105**	**.243**	**.147**
age18_29	**.280**	**.158**	**-.448**	**.224**
age30_49	.096	.127	-.228	0 .179
peop3_4	0.039	0.558	-0.012	0.914
peop_5above	**0.106**	**0.018**	0.012	0.853
budget	**.165**	**.049**	**.217**	**.068**
color	**.108**	**.043**	0.077	0.601
size	.087	.056	-0.036	0.798
fresh	**.224**	**.064**	-.429	.190
price	**-.122**	**.050**	.067	.070
availability	**.121**	**.054**	.045	.077
constant	**-1.144**	**.395**	.344	.545
Rho (Covariance Parameter)	1.165293	0.0875	**1.503**	**.066**
Model fit (P-value)	0.000		0.000	
Number of Observations	586		548	