FACTORS TO BE CONSIDERED IN CHOOSING THE TYPE OF CONVENTIONAL DWELLING: MONGOLIAN CASE

Myagmarsuren Boldbaatar, PhD candidate,
Department of Economics and Statistics, College of Business and Economics,
Korea University, Republic of Korea

Abstract

This study attempted to determine the factors that influence the choice of the type of dwelling using the sample from the "Urban Poverty and in-Migration" survey, which was conducted by Population Teaching and Research Center at School of Economic Studies, National University of Mongolia in 2004. Regression models for categorical data such as Ordered Logit and Multinominal Logit models were utilized to answer the research question under consideration. Due to the data availability, ten factors were used in this analysis but during the estimation stage two factors were excluded from the analysis on account of the statistical significance. It concludes that the factors, namely age, sex and education level of the head of family, number of person in household, percentage of migrants in household, number of children in household, total revenue of household per year and percentage of food expenses in total expenditure per year have significant effect on probability to live in particular type of dwelling. And their effect directions were exactly same as the common sense.

KEY WORDS: Choice of dwelling, Ordered Logit Model, Multinominal Logit Model

INTRODUCTION

Mongolia is showing progress in overall economic situation, with economic growth averaging 6.1 per cent for the recent years. Nowadays 40.7 per cent of population, while 38.1 per cent of households are living in Ulaanbaatar (UB)- the capital city of Mongolia[34].

Although, economic situation is improving, the living conditions of the citizens of Ulaanbaatar have not improved. They are still living in dwelling which have no basic living facilities such as centralized drinking water system, indoor flush toilet and centralized sewage water system. According to the data of Statistical Office of Ulaanbaatar in 2009, the 39.2 per cent of UB households are living in *Apartment*, 33.8 per cent of households are living in *House*, 26.4 percent of households are living in *Ger* and 0.6 per cent of households are homeless.

I should clarify here the understandings of the words - *Apartment, House* and *Ger* in Mongolian case. *"Apartment"* has same interpretation as other nations but the *"House"* has little bit different. Generally, *Houses* are 1-2 floored and have from two[35] till many rooms; the only difference is there are no basic living facilities. The *Ger* is Mongolian traditional dwelling which is best for nomads, easy to move, it takes no more than two hours to pack and put up. If I would describe the *Ger* as a house, it is like single-circle roomed, moveable house.

I am studying the urban area which is making to lose the biggest advantage of *Ger*. Thus I can sort the types of dwelling considering in living conditions as *Ger, House* and *Apartment* in ascending order.

CONCEPTUAL FRAMEWORK

Taking the data availability into consideration, I would use intuitive approach mostly to find out the factors that we are searching for. For some case, it is convenient to use simple analyzing method such as correlation.

First, family income would be key factor of the study. If the family has more income, they would have more chance to build a new house or to buy an apartment.

[34] NSO, 2010. Mongolia Statistical Yearbook-2009, National Statistical Office of Mongolia, Ulaanbaatar

[35] Including kitchen

Second, family income is from the employment of the household members. If they have no work, there would be no income as well as they will be live in poor condition. "There are strong linkages between the employment of the head of a household and that household's poverty status. Working with data from the 2002-2003 Living Standards Measurement Survey, the report calculates a "poverty likelihood ratio" (PLR) for households, the ratio of each group's poverty incidence to that of the overall population sorted according to whether the head is working and the sort of work. Households headed by an individual who is of working age but is economically inactive – i.e. neither working nor looking for work – are the most likely to be poor, with a PLR of 1.42, notably more likely even than households headed by an unemployed individual, who have a PLR of 1.33."[36] Thus I am choosing the fact, which is head of the household works or not, as a factor. In addition, it is reasonable to choose the percentage of working person in household as a factor as well.

Third, some individual characteristic of the head of household such as age, sex and attained highest education level could be useful in this analysis. The reason why I am choosing is particular households' general status depends on the head of family according to the social customs of Mongolia. He or she, the head of household, always takes the principal decision regarding to the household. Furthermore, age of the head of household can represent the age of family whether they are young or old. Initially, young families don't have wealth but as time goes by they would accumulate wealth and would live in a better living condition. For the education, I have studied the correlation between education level and the type of dwelling and the result showed significant association[37].

Fourth, the internal migration tends to the move to the urban area in order to find job as well as to move closer to services for education and health during the last decade. They move to the urban area due to dzud[38] which makes many herders to lose their livestock. Mainly, migrant household lives in a *Ger* and is expanding the *Ger* district of Ulaanbaatar. Thus I decided to choose a variable that could stand for migration status of household. According to the data availability, there are two type of variable, one of them is the head of household is migrant or not, the other one is percentage of migrants in household.

Fifth, the number of total person and children in the household could be considered as a factor. As number of person in household increases, the opportunity of access to the better living condition could be limited if the additional person does not earn the income.

Finally, I have noted before about the importance of household income. Now it is time to consider the usage of household income. The percentage of food expenses in total expenditure could be reasonable explanatory variable in our analysis. The income should first cover all the food expenses and the surplus would be used to finance non food and service expenses. The point is that the chosen variable can stand for the sufficiency of household income. If the percentage of food expenses is high then the opportunity of improving living condition would be limited.

ECONOMETRIC MODELS AND ESTIMATION

Model specification of the study

Since the type of conventional dwelling is the categorical data, it can be viewed as ordinal or nominal variables. Thus I would employ the ordered logit model as well as multinominal logit model in order to find out appropriate regression model.

Ordinal Logit Model

$$usetype = \begin{cases} 1 \Rightarrow Ger \;\; if \;\; -\infty \leq \\ 2 \Rightarrow House \; if \; \tau_1 \leq \\ 3 \Rightarrow Apartment \;\; if \tau_2 \end{cases}$$

[36] Government of Mongolia, UNDP, 2008. Mongolia Human Development Report 2007: Employment and Poverty, (pp. 13-15) 2007, Ulaanbaatar

[37] Refer to the Table 1 in the appendix.

[38] A **zud** or **dzud** (Mongolian: зуд) is a Mongolian term for an extremely snowy winter in which livestock are unable to find fodder through the snow cover, and large numbers of animals die due to starvation and the cold. (Wikipedia)

$$y^* = \beta_1 + \beta_2 age + \beta_3 sex + \beta_4 educ1 + \beta_5 educ2 + \beta_6 educ3 + \beta_7 educ4 + \beta_8 totper + \beta_9 mig + \beta_{10} mig_{percent} + \beta_{11} work + \beta_{12} work_{percent} + \beta_{13} childnum + \beta_{14} totrev + \beta_{15} food + \varepsilon$$

Here we assume that $\varepsilon \sim logistic$ then the model becomes Ordered Logit Model.

$Prob(housetype_i = 1|x_i) = F(\tau_m - \beta_i x_i) - F(\tau_{m-1} - \beta_i x_i)$

Here, $\mathrm{F}(\varepsilon) = \Lambda(\varepsilon) = \frac{\exp(\varepsilon)}{1+\exp(\varepsilon)}$

The explanations of the chosen variables are shown in Table 1.

Table 1: **Variable description**

y^*	Latent variable/dependent variable
Independent variables	
age	age of the head of family, in years
sex	sex of the head of family, 0 if female, 1 if male
educ1	= 1 if education level is secondary, 0 otherwise
educ2	= 1 if education level is primary, 0 otherwise
educ3	= 1 if education level is college/graduated, 0 otherwise
educ4	= 1 if education level is vocational, 0 otherwise
totper	number of person in household
mig_percent	percentage of migrants in household, %
work	=1 if he/she works, 0 otherwise
work_percent	percentage of working persons in household, %
childnum	number of children in household
totrev	total revenue of household per year, thousand tugrugs
food	percentage of food expenses in total expenditure per year, %

<u>*Multinominal Logit Model*</u>

We can capture the effects of the chosen explanatory variables by estimating two binary logit models,

$$Ln\left[\frac{\Pr(G|x)}{\Pr(A|x)}\right] = \beta_{1,G|A} + \sum \beta_j x_j$$

$$Ln\left[\frac{\Pr(H|x)}{\Pr(A|x)}\right] = \beta_{1,H|A} + \sum \beta_j x_j$$

were,x_j-chosen factors, $\Omega_{\mathrm{H|A}}(x) = \frac{\Pr(H|x)}{\Pr(A|x)}$ and $\Omega_{\mathrm{G|A}}(x) = \frac{\Pr(G|x)}{\Pr(A|x)}$ are the odds ratio,

$$Prob(housetype = m|x_i) = \frac{\exp(\beta_m x_i)}{\sum_{k=1}^{J} \exp(\beta_k x_i)}$$

The models were estimated by STATA package.

Data source of the study

The data source is from the "Urban Poverty and in-Migration" survey (2004), which cover 6847 residents in 1500 households in Ulaanbaatar, capital city of Mongolia. This Survey was

conducted by Population Teaching and Research Center, at School of Economic Studies, National University of Mongolia. The data set was collected through individual interviews of each household.

The summary statistics of the variables are shown in Table 2.

Table 2: **Summary Statistics**

Variable	Mean	Std. Dev.	Min	Max
totper	4.6	2.0	1	15
age	46.0	14.5	17	90
childnum	2.0	1.5	0	10
totrev	3037.8	2036.1	130.8	1618.6
food	38.1	16.2	7.99	95.28
mig_percent	9.4	23.3	0	100
work_percent	33.4	23.2	0	100

Estimation

Ordered Logit Model (OLM)

First, I estimated the regression taking the all chosen variable as a explanatory factor (Model 1) then if some of the regression coefficient are been insignificant, excluding that factor from the model and re-estimated the model (Model 2). The results are shown in Table 3.

Table 3: **OLM Estimation Result. Dependent variable: Type of dwelling**

Independent Variables	Model 1	Model 2
totper	-0.0783 (0.040)	-0.0823 (0.040)
mig_percent	-0.0272 (0.003)	-0.0273 (0.003)
age	0.0201 (0.005)	0.0191 (0.004)
sex	-0.3417 (0.140)	-0.3352 (0.139)
educ1	**0.3361 (0.407)**	-
educ2	0.7443 (0.380)	0.4709 (0.204)
educ3	1.0540 (0.391)	0.7902 (0.223)
educ4	2.3010 (0.397)	2.0454 (0.230)
work	**0.0138 (0.152)**	-
work_percent	**0.0022 (0.003)**	-
childnum	-0.2624 (0.056)	-0.2599 (0.055)
totrev	0.0003 (0.000)	0.0003 (0.000)
food	-0.0087 (0.005)	-0.0088 (0.005)
LR chi-square	599.93	598.36
Prob > Chi-square	0.0000	0.0000
Pseudo R-square	0.2045	0.2040
Observations	1407	1407

Note: () standard error. Insignificant variables are shown as bold.

Before excluding the three variables (*educ1, work, work_percent*) the LR test was employed. Table 4 illustrates those three variables have no effects on the type of dwelling. Education level being primary and non educated have same effects on choosing type of the conventional dwelling. The percentage of working persons in household has no effect on categorical dependent variable as well. Perhaps, it is due to its skewed distribution, the 83 per cent of total household is non migrant; only 2 per cent of total household is migrant. Using this kind of data, it would be difficult to catch the significant effect of this variable. My expectation about the work variable was reasonably high but the result was not good enough. Maybe, this variable is not suitable with the chosen model which is requiring further study.

Table 4: **Likelihood-ratio test**

(Assumption: tested in LRTEST_0)	LR chi-square (3)=1.57 Prob > chi-square = 0.6661

The Model 2 has pretty good results, all Z and LR chi square statistics are statistically significant. But it is important to check the parallel regression assumption before using this model.

Table 5: **Parallel regression assumption test results**

Test type	Chi-square	p>chi-square	df
Approximate likelihood-ratio test	50.36	0.000	10
Brant test	56.28	0.000	10

From the Table 5, both tests show that the parallel regression assumption can be rejected at the 0.01 level. In this case, as a noted by Long and Jeremy when the assumption is rejected, alternative models should be considered that do not impose the constraint of parallel regression.

Multinominal Logistic Model Estimation (MNLM)

Since we cannot use the ordinal logit model, let's concentrate on the MNLM. The result is shown in Table 6.

Table 6: MNLM Estimation Result. Dependent variable: Type of the dwelling				
Independent Variable	**Ger**		**House**	
	Coef. (Std. Err.)	**P value**	**Coef. (Std. Err.)**	**P value**
Totper	0.126 (0.064)	0.05	0.108 (0.052)	0.039
mig_percent	0.037 (0.004)	0.00	0.010 (0.004)	0.017
Age	-0.028 (0.007)	0.00	-0.021 (0.006)	0.00
Sex	0.430 (0.222)	0.053	0.319 (0.172)	0.064
educ2	-0.787 (0.326)	0.016	-0.395 (0.277)	0.154
educ3	-1.186 (0.352)	0.001	-1.023 (0.295)	0.001
educ4	-2.668 (0.386)	0.00	-2.043 (0.293)	0.00
Childnum	0.437 (0.090)	0.00	0.384 (0.073)	0.00
Totrev	0.000 (0.000)	0.00	-0.0003 (0.0001)	0.00
Food	0.016 (0.008)	0.037	0.017 (0.006)	0.007
Intercept	0.271 (0.699)	0.698	0.515 (0.547)	0.347
Pseudo R-square	0.2235			
LR chi-square (20)	655.67			
Prob > Chi-square	0.0000			
Observations	1407			

From the Table 6, it is easy to see that all Z and LR statistics are statistically significant except *constant terms* and *educ2* in case of dwelling.

For testing for independent variables, Wald test is tighter than LR test.

Table 7 illustrates that all the coefficients associated with the independent variables (except *totperson* and *sex)* are statistically significant at 0.05 level.

Table 7: Wald test result			
Independent Variables	**Chi-square**	**df**	**P>Chi-square**
totper	5.2	2	0.074
mig_percent	115.5	2	0.000
age	20.1	2	0.000
sex	4.8	2	0.089
educ2	5.8	2	0.05
educ3	14.8	2	0.001
educ4	65.4	2	0.000
childnum	32.4	2	0.000
totrev	45.9	2	0.000
food	7.9	2	0.02

Ho: All coefficients associated with given variable(s) are 0.

We can reject the hypothesis that *Ger* and House, *Ger* and Apartment, House and Apartment are indistinguishable (Table 8).

Table 8: Wald tests for combining alternatives			
Alternatives tested	**Chi-square**	**df**	**P>chi-square**
Ger - House	88.9	10	0.000
Ger - Apartment	259.2	10	0.000
House - Apartment	254.8	10	0.000

Ho: All coefficients except intercepts associated with a given pair of alternatives are 0 (i.e., alternatives can be combined).

The important test for MNLM is independence of irrelevant alternatives (IIA). Here, I applied Small and Hsiao test of IIA (Table 9).

Table 9: Small-Hsiao tests of IIA assumption						
Omitted	**lnL(full)**	**lnL(omit)**	**Chi-square**	**df**	**P>chi-square**	**Evidence**
Ger	-301.849	-298.031	7.635	11	0.746	for Ho
House	-174.663	-170.936	7.455	11	0.761	for Ho

Ho: Odds (Outcome-J vs Outcome-K) are independent of other alternatives.

In our analysis, each test indicates that IIA assumption has not been violated.
We have tested the MNLM by many tests and the results were all statistically significant. Thus we can apply the estimated model hereafter.

All types of R-squares are relatively high in this model (Table 10).

Table 10: **Measures of Fit for MNLM of type of dwelling**

Log-Likelyhood Intercept Only:	-1466.5	Log-Likelyhood Full Model:	-1138.7
D(1385):	2277.4	LR(20):	655.7
		Prob > LR:	0.00
McFadden's R-square:	0.22	McFadden's Adj R-square:	0.21
ML (Cox-Snell) R-square:	0.37	Cragg-Uhler(Nagelkerke) R-square:	0.43
Count R-square:	0.64	Adj Count R-square:	0.36

RESULTS OF THE STUDY

What we have found out so far is the variables used in MNLM have significant effect on the type of dwelling. In addition, model statistics are all statistically significant.

Table 11 shows the estimates of discrete change from our model for type of dwelling. For the *totperson* variable; each additional person in the household, the probability of living in *Ger* increases by 0.9 per cent and increases the probability of living in House by 1.8 per cent while decreases the probability of living in Apartment by 2.8 per cent, holding all other variables at their means.

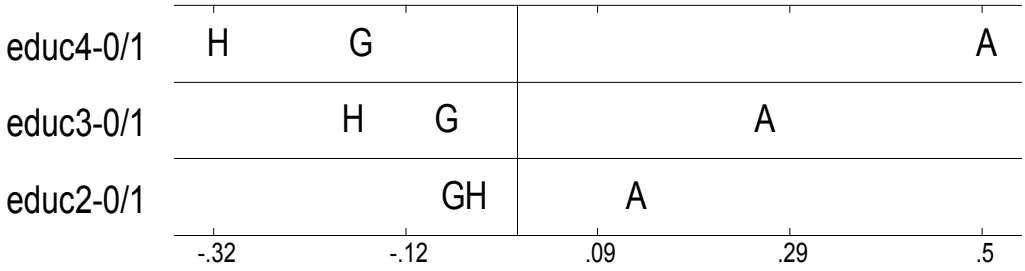
Table 11: **Discrete Change in Probability for a MNLM of type of dwelling**

Variable	Change	$\bar{\Delta}$	*Ger*	House	Apartment
totperson	ΔRange	0.229	0.127	0.216	-0.344
	Δ1	0.018	0.009	0.018	-0.028
	Δσ	0.038	0.019	0.038	-0.057
mig_percent	ΔRange	0.417	0.625	-0.231	-0.394
	Δ1	0.003	0.004	0.0001	-0.004
	Δσ	0.064	0.094	0.002	-0.095
age	ΔRange	0.262	-0.148	-0.246	0.394
	Δ1	0.004	-0.002	-0.003	0.006
	Δσ	0.054	-0.032	-0.050	0.081
sex	0→1	0.057	0.033	0.053	-0.086
educ2	0→1	0.081	-0.073	-0.049	0.122
educ3	0→1	0.173	-0.080	-0.179	0.259
educ4	0→1	0.330	-0.171	-0.324	0.495
childnum	ΔRange	0.400	0.250	0.350	-0.600
	Δ1	0.065	0.031	0.066	-0.098
	Δσ	0.095	0.046	0.097	-0.143
totrev	ΔRange	0.521	-0.306	-0.474	0.781
	Δ1	0.00006	-0.00004	-0.00004	0.00008
	Δσ	0.114	-0.090	-0.081	0.171
food	ΔRange	0.219	0.081	0.248	-0.328
	Δ1	0.003	0.001	0.003	-0.004
	Δσ	0.044	0.016	0.049	-0.066

Note: 0→1 is change from 0 to 1; Δ1 is centered change of 1 around the mean; Δσ is centered change of 1 standard deviation around the mean; ΔRange is change from the minimum to its maximum. $\bar{\Delta}$ is the average absolute discrete change. All the variables are held at their mean.

For the *education* variable; if a head of household has college degree, his household probability of living in Apartment is 49.5 per cent greater than a head of household whose education level is less than primary, holding all other variables at their means. Summing up all the results relative to the education variables, it shows the following pattern. As education level increases, the average absolute discrete change increases. It tells us that as education level of head of household increases, it is likely to move from *Ger* to Apartment or from House to Apartment or from *Ger* to House (Figure 1).

Figure 1: **Change in Predicted Probability for type of dwelling (0→1)**

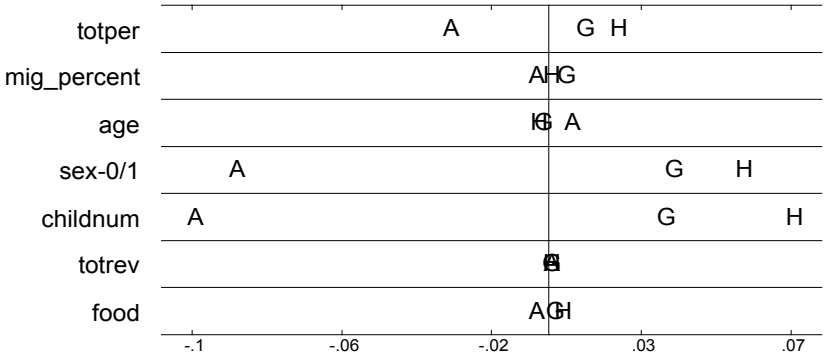


Effect directions of the variables for the probabilities of living in *Ger* and House are exactly same. The opposite directions of the above are shown for the probabilities of living in Apartment. Although the effect directions of the variables for the probabilities of living in *Ger* and House are exactly same but they are differed in size of effect. The size of effects of the variables for the probability of living in House are greater than the size of effects of the variables for the probability of living in *Ger* (refer to the Figure 2 and Table 12).

Table 12: **Effect directions of the Probability for MNLM of Type of dwelling**

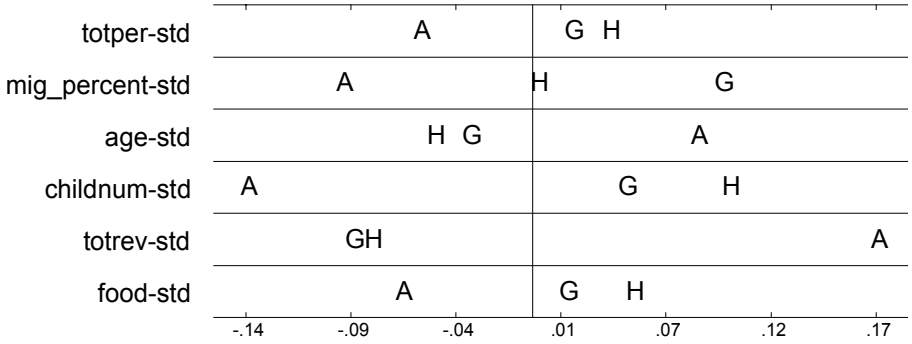
Variable	*Ger*	House	Apartment
totperson	+	+	−
mig_percent	+	+	−
age	−	−	+
sex	+	+	−
educ	−	−	+
childnum	+	+	−
totrev	−	−	+
food	+	+	−

Figure 2: **Change in Predicted Probability for type of dwelling (Δ1)**



For the dummy variable *sex*: holding all other variables at their means, being male increases the both probability of living in *Ger* and House by 3.3 per cent and 5.3 per cent, respectively and decreases the probability of living in apartment by 8.6 per cent.

Figure 3: **Change in Predicted Probability for type of dwelling (Δσ)**



It is easy to see from the Figure 3, the average absolute change for a standard deviation change in *totrev* is 11.4 per cent, the maximum value from the all of average absolute change. Thus *household income* contributes more relatively to the explanation of the other (numerical) variables. The effect of *totrev* is largest on the probability of living in apartment, where the expected change for a standard deviation in *totrev* is 17.1 per cent. Furthermore, the average absolute change for a standard deviation change in *childnum* is -9.5 per cent, the minimum value from the all of average absolute change. The effect of *childnum* is smallest on the probability of living in apartment, where the expected change for a standard deviation in *childnum* is -14.3 per cent.

CONCLUSIONS

This study has attempted to determine the factors which influence to live in a particular type of dwelling. The sample is from the Urban Poverty and in-Migration survey in 2004, which covers 1500 households in Ulaanbaatar. Ten plausible factors were used in analysis but during the estimation stage two of them are excluded. Using both ordered logit and multinominal logit models, significant ML estimates and tests were reported.

The factors, namely age, sex and education level of the head of family, number of person in household, percentage of migrants in household, number of children in household, total revenue of household per year and percentage of food expenses in total expenditure per year have significant effect on probability to live in particular type of dwelling. The excluded two factors have strong logical statements to be considered as a factor in this type of analysis. Therefore, it leaves us further study.

REFERENCE

Damodar N. Gujarati, D. C. (2009). Chapter 15. In D. C. Damodar N. Gujarati, *Basic Econometrics* (pp. 541-591). Singapore: McGraw-Hill.

Greene, W. H. (2003). Chapter 21. In W. H. Greene, *Econometric Analysis* (pp. 719-729). New Jersey: Prentice Hall.

Government of Mongolia, UNDP, (2007). *Mongolia Human Development Report 2007: Employment and Poverty.* Ulaanbaatar: Admon LLC

J.Scot Long, J. F. (2001). *Regression Models for Categorical Dependent Variables Using Stata.* Texas: A Stata Press Publication.

Long, J. S. (1997). *Regression Models for Categorical and Limited Dependent Variables.* California: Sage Publications, Inc.

NSO, 2000-2010. *Mongolian Statistical Yearbook.* (2000-2009). Ulaanbaatar: National Statictical Office of Mongolia, Ulaanbaatar.

Statistical Office of Ulaanbaatar City. (n.d.). Retrieved 2010 оны 12 22 from Website of Statistical Office of Ulaanbaatar City: http://www.statis.ub.gov.mn/

Wikipedia. (n.d.). Retrieved 12 22, 2010, from Wikipedia, the free encyclopedia: http://en.wikipedia.org/wiki/Zud

Wooldridge, J. M. (2009). Chapter 17, Chapter 19. In J. M. Wooldridge, *Introductory Econometrics: A Modern Approach* (pp. 575-612, 668-687). South-Western Cengage Learning.