ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 228, 2009

*Joanna Landmesser**

THE SURVEY OF ECONOMIC ACTIVITY OF PEOPLE IN RURAL AREAS – THE ANALYSIS USING THE ECONOMETRIC HAZARD MODELS

Abstract. This paper discusses the duration of employment periods in the rural population. Estimating the risk models, the direct risk of leaving the job is calculated. We estimate hazard rate models to assess the effect of such factors as: gender, age, education level, place of residence and employment status, on the individual's employment duration. We establish differences between economic activity of people in rural and urban areas and those between residents of rural areas from various polish provinces. To estimate, we use data from the Labour Force Survey in Poland.

Key words: economic activity of people, hazard models, Cox model.

I. INTRODUCTION

The aim of this paper is to analyse the duration of employment periods in the rural population, using the survival analysis. The hazard model can be a suitable tool for the analysis of the employment period duration. Estimating the non-competing risk models and competing risk models we calculate the conditional probabilities of transition from the employment state to the unemployment state or out-of-the-labour-force state (e.g. retirement, disability, studies). Hazard models comprise not only present employment duration as a determinant for the probability of leaving the job, but also other observable characteristics of individuals such as gender, age, education level, employment status (own-account workers or paid employees), ownership sector of the job (public or private sector). We try to explore differences between economic activity of people in rural and urban areas.

* Ph.D., Chair of Econometrics and Statistics, Warsaw University of Life Sciences.

This paper is written as research project "The Survey of Employment Activity of People in Rural Areas Using the Survival Analysis" (SGGW grant 504-08270017).

II. THE ANALYSIS METHOD

Statistical duration models are increasingly used by econometricians to analyse various economic problems. The first intensive application of these models is the analysis of individual unemployment duration data by Lancaster (1979). Econometric models of durations are models of the length of time spent in a given state before transition to another state. The duration analysis is also called the survival analysis, the failure time analysis or the hazard analysis (for wider overview see Kalbfleisch, Prentice (1980), Cox, Oakes (1984), Kiefer (1988), Hosmer, Lemeshow (1999), Cameron, Trivedi (2005)).

Let T be a nonnegative random variable describing duration in any state. The distribution function of T is denoted F and is defined as $F(t) = \Pr[T \leq t]$. The density function of the duration variable T is $f(t) = dF(t)/\mathrm{dt}$. There may be provided an additional function called the survival function $S(t)$ which gives the probability that the duration T is greater than t (the probability of surviving past t), $S(t) = \Pr[T > t] = 1 - F(t)$. The most frequently applied demonstration of the duration period distribution is hazard function $h(t)$. It is the limit of probability that the spell is completed during the interval $[t, t+dt]$ given that it has not been completed before the time t, for $dt \rightarrow 0$.

$$h(t) = \frac{f(t)}{S(t)} = \lim_{dt \to 0} \frac{\Pr[t \leq T < t + dt \mid T \geq t]}{dt} \tag{1}$$

The hazard rates describe the intensity of transition from one state to another.

Hazard models usually comprise present duration of the phenomenon as a determinant for the probability of its occurrence and also other parameters. In the proportional hazard models, the conditional hazard rate $h(t|X)$ can be factored into separate functions: $h(t|X) = h_0(t)g_0(X) = h_0(t)\exp(X\beta)$, where $h_0(t)$ is called the baseline hazard and $exp(X\beta)$ is a function of explanatory variables vector X. The characteristics of hazard function change proportionally to the influence of explanatory variables.

Cox's (1972) approach to the proportional hazard model is the semiparametric method of analysing the effect of covariates on the hazard rate. The Cox model states that the hazard rate for the jth subject in the data is $h(t|X_j) = h_0(t)\exp(X_j\beta)$. Compared with the parametric approaches, the advantage of the semiparametric Cox model is that we have no need to make assumptions about baseline hazard; $h_0(t)$ is left unestimated. One subject's hazard

is a multiplicative replica of another's. Comparing subject i to subject j, the model states that

$$\frac{h(t|X_i)}{h(t|X_j)} = \frac{h_0(t)\exp(X_i\beta)}{h_0(t)\exp(X_j\beta)} = \exp\left[(X_i - X_j)\beta\right], \tag{2}$$

which is constant assuming that the covariates X_i and X_j do not change over time. The parameters of Cox hazard models β can be estimated by partial likelihood method.

In many situations, there are several possible risks of failure. The competing risks model formulation is applicable to modelling time in one state when the exit occurs to a number of competing states.

III. THE SUBJECT OF THE RESEARCH

To find differences between economic activity of people in rural and urban areas we use data from the Labour Force Survey in Poland (Badanie Aktywności Ekonomicznej Ludności Polski BAEL). The survey concentrates on the situation of population from the point of view of economic activity of people, i.e. the fact of being employed, unemployed or economically inactive in the reference week. Economically active population includes all persons aged 15 and over who are considered as employed as well as unemployed persons. On the basis of the retrospective questions in BAEL-questionnaire we can conclude how long one is employed or how long one was employed in the previous job and whether the exit was to unemployment state or to out-of-the-labour-force state (OLF-state).

The whole BAEL-sample has been limited to a subsample of 26038 persons aged 18 years or older, who in the period 1994-2002 are or were employed at least for one year. At the end of the study, these people either were full-time or part-time employed (18001 persons) or were classified as unemployed (2630) or out-of-labour-force (5407) (economically inactive). 11178 reside villages and 14860 towns and cities. Among village residents 4846 support families from agriculture and 6332 support families from different sectors. The individual employment duration in the previous job (in years) for each person built a variable *duration*. The explanatory variables for the hazard models are presented in Table 1.

Table 1. Definitions of Variables

Variable	Description
male	1 if individual is male
age1	1 if individual is 18 years or older but younger than 24
age2	1 if individual is 25 years or older but younger than 34
age3	1 if individual is 35 years or older but younger than 44
age4	1 if individual is 45 years or older but younger than 54
age5	1 if individual is 55 years or older
edu1	1 if individual has tertiary education level
edu2	1 if individual has vocational secondary education level
edu3	1 if individual has general secondary education level
edu4	1 if individual has basic vocational education level
edu5	1 if individual has lower secondary or primary education level
edu6	1 if individual has incomplete primary education level
ruralarea	1 if the place of residence is rural area
partner	1 if individual is married
headhh	1 if individual is the head of the household
farm	1 if individual works in the private farm
agriculture	1 if individual works in agricultural sector
private	1 if individual works in the private sector
employer	1 if individual is employer (hires one or more employees)
own-account	1 if individual is own-account worker and hires no employees
employees	1 if individual is paid employee
familyworker	1 if individual is contributing family worker
disabled	1 if individual is disabled

IV. RESULTS OF HAZARD MODELS ESTIMATION

Significant differences persist between employment durations in rural and urban areas. The activity of people in rural areas is different from that in urban areas. The results of the econometric analysis for the timing of exit from the job are reported in Table 2.

Table 2. Results of Cox models estimation for the risk of leaving the job in rural and urban areas

Variable	Non-competing hazard				Competing hazard							
	job→another state				job→unemployment				job→OLF			
	rural		urban		rural		urban		rural		urban	
	Hazard Ratio		Hazard Ratio		Hazard Ratio		Hazard Ratio		Hazard Ratio		Hazard Ratio	
1	2	3	4	5	6	7	8	9	10	11	12	13
male	0.760	***	0.846	***	*1.402*	***	*1,381*	***	*0,542*	***	*0,596*	***
age1	3.660	***	3.924	***	2.076	***	1,906	***	*5,953*	***	*7,401*	***
age2	2.255	***	2.018	***	1.449	***	1,431	***	3,135	***	2,615	***
age4	0.821	***	0.760	***	0.774	***	0,725	***	1,137		1,004	

Table 1 (cont.)

1	2	3	4	5	6	7	8	9	10	11	12	13
age5	0.794	***	0.808	***	0.167	***	0,275	***	1,435	***	1,448	***
edu2	*1.795*	***	1.472	***	1.971	***	*1,887*	***	1,689	***	*1,347*	***
edu3	*1.988*	***	1.747	***	2.599	***	*2,547*	***	1,724	***	*1,505*	***
edu4	*3.720*	***	2.353	***	3.989	***	*4,085*	***	3,508	***	*1,832*	**
edu5	*2.905*	***	2.418	***	2.786	***	*3,376*	***	2,909	***	*2,157*	***
edu6	*2.191*	***	1.393		1.375		*2,883*		*2,135*	***	*1,547*	
partner	0.912	**	0.757	***	0.778	***	0,633	***	1,135	**	0,948	
headhh	0.507	***	0.514	***	0.312	***	0,343	***	0,664	***	0,661	***
farm	*0.099*	***	0.120	***	0.027	***	0,078	***	0,130	***	0,142	***
agriculture	*1.159*	*	1.095		1.423	***	1,196		1,067		1,098	
private	*1.529*	***	*1.610*	***	1.873	***	2,359	***	1,312	***	1,367	***
own-account	0.875		1.380	***	1.320		1,659	***	0,755		1,213	
employees	1.487	**	1.668	***	1.888	**	1,788	***	1,396	*	1,650	***
familyworker	0.339	***	0.721		1.058		0,843		0,296	***	0,623	*
disabled	2.412	***	2.292	***	0.423	***	0,632	***	2,999	***	3,265	***
kujaw.-pom.	0.895		0.841	**	0.743	*	0,763	**	1,060		0,940	
lubelskie	*0.564*	***	0.877		0.551	***	0,894		0,634	***	0,894	
lubuskie	0.948		0.892		0.822		0,713	***	1,068		1,041	
łódzkie	0.857		0.918		0.726	*	0,719	***	1,001		1,116	
małopolskie	*0.572*	***	0.617	***	0.514	***	0,465	***	0,615	***	0,714	***
mazowieckie	0.851		0.966		0.578	***	0,713	***	1,087		1,232	**
opolskie	0.888		0.740	***	0.626	**	0,436	***	1,108		1,016	
podkarpackie	*0.555*	***	0.872		0.561	***	0,715	**	0,581	***	1,031	
podlaskie	0.837		1.226	**	0.732		1,014		0,957		1,386	***
pomorskie	1.055		0.861	**	0.959		0,586	***	1,135		1,122	
śląskie	0.805	**	0.892	*	0.536	***	0,584	***	1,007		1,158	*
świętokrzyskie	0.787	**	1.133		0.464	***	0,722	**	1,011		1,471	***
warmińskie	*1.199*	*	1.121		1.145		0,822		1,245		1,364	***
wielkopolskie	0.794	**	0.801	***	0.606	***	0,579	***	0,952		0,985	
zach.pomor.	*1.359*	***	1.072		1.210		0,856		1,469	***	1,276	***
obs.no.	11178		14860		11178		14860		11178		14860	
lnL	-23485.98		-40440.71		–7443.88		–14489.78		–15722.71		–25449.47	

***; **; * – significant at 1%, 5%, 10% level respectively.
Source: own computations.

The hazard of being unemployed in the case of a man is bigger than in the case of a woman (40,2% bigger in rural area, 38,1% in urban area). But the hazard of being economically inactive is smaller in the former case (47,8% and 40,4% respectively) certainly due to traditional role of a woman as a housekeeper.

The young people aged 18-24 in urban areas are much more exposed to being out-of-labour-force than young people in rural areas, which may be influenced by the greater access to education in towns and cities.

The effect of increasing hazard in losing the job as a result of lower education level is stronger in rural areas. Also there it is more probable that the lack of any level of education will lead to the OLF-state than to being unemployed. About 30% of agricultural work force has only basic vocational education and, in consequence, it is very difficult for these people to obtain jobs. In towns and cities the lower education level than the tertiary one does not easily eliminate people from the labour force, but increases the risk of unemployment.

The employment in the private agricultural farm decreases the hazard of quitting the job by 80,1%. But many inhabitants of rural areas are not farm holders and were usually previously employed in the state-owned agricultural farms. However, a large number of these farms went bankrupt after the political system change in 1989. The former employees of these farms turned out to be the most passive and helpless social group in Poland. Even if they worked later in agricultural sector, they are still exposed to 15,9% risk of job loss. The highest hazard rate of exiting job in rural areas was observed in those polish provinces where the former state-owned farms had been located: Zachodniopomorskie and Warmińsko-Mazurskie, whereas the lowest rate characterised such provinces as: Lubelskie, Małopolskie and Podkarpackie.

If an individual works in the private sector the risk of leaving the present job is bigger in rural areas as well as in urban areas.

Among inhabitants of rural areas employed in agriculture younger age increases the risk of being passive (see Table 3). However, as the time passes this risk decreases faster than in the case of those working out of agriculture. The lack of education strongly disturbs in different sectors than agricultural. The job in the private firm protects against OLF-state in the agricultural sector, but increases the OLF-risk behind this sector. It is the paid employees who are mostly exposed to OLF in agriculture.

Table 3. Results of Cox models estimation for the transition into OLF-state among village residents supporting families from agricultural and different from agricultural sector

Variable	Agricultural activity		Out of agricultural activity		Variable	Agricultural activity		Out of agricultural activity	
	Hazard Ratio		Hazard Ratio			Hazard Ratio		Hazard Ratio	
1	2	3	4	5	6	7	8	9	10
male	0.649	***	0.472	***	kujaw.-pom.	0.956		0,843	
age1	*28.686*	***	*5.190*	***	lubelskie	0.525	***	0,562	***
age2	*2.408*	***	*3.413*	***	lubuskie	1.020		0,971	
age4	*0.762*		*1.130*		łódzkie	0.764		0,969	

Table 3 (cont.)

1	2	3	4	5	6	7	8	9	10
age5	*0.623*	**	*1.658*	***	małopolskie	0.398	***	0,755	**
edu2	1.794		1.619	***	mazowieckie	0.958		0,939	
edu3	2.555	*	1.544	**	opolskie	0.882		1,117	
edu4	3.647		3.552	***	podkarpackie	0.415	***	0,639	***
edu5	2.142	*	3.019	***	podlaskie	0.780		0,969	
edu6	1.606		*5.587*	***	pomorskie	1.186		1,052	
partner	0.990		1.188	**	śląskie	1.724	*	0,832	
headhh	0.601	***	0.660	***	świętokrzyskie	0.869		0,955	
private	*0.733*	*	*1.306*	***	warmińskie	1.282		1,167	
own-account	1.050		0.944		wielkopolskie	1.015		0,792	
employees	*7.231*	***	1.274		zach.pomor.	1.547	*	1,274	
familyworker	0.388	***	0.514		*obs.no.*	4846		6332	
disabled	2.256	***	4.277	***	*lnL*	–5694.20		–8870.28	

***; **; * – significant at 1%, 5%, 10% level respectively.
Source: own computations.

V. CONCLUSIONS

To conclude, hazard models are proper tools to the analysis of the people's economic activity. Estimating the hazard models for various subsamples, e.g. rural and urban areas residents, enables making comparisons concerning the disposition of people to change the activity states. When the estimation of the hazard models with one exit provides no real description about the impact of explanatory variables and the duration on the process analysed, the competing risk hazard models should be used.

REFERENCES

Cameron A.C., Trivedi P.K. (2005), *Microeconometrics: Methods and Applications*, Cambridge University Press, New York.

Cox D.R. (1972), Regression Models and Life Tables (with Discussion), *Journal of the Royal Statistical Society*, Series B 34, p. 187–220.

Cox D.R., Oakes D. (1984), *Analysis of Survival Data*, Chapman and Hall, London.

Hosmer D., Lemeshow S. (1999), *Applied Survival Analysis: Regression Modeling of Time to Event Data*, John Wiley and Sons, New York.

Kalbfleisch J., Prentice R. (1980), *The Statistical Analysis of Failure Time Data*, John Wiley and Sons, New York.

Kiefer N. (1988), Economic Duration Data and Hazard Functions, *Journal of Economic Literature*, 26, p. 646–679.

Lancaster T. (1979), Econometric Methods for the Duration of Unemployment, *Econometrica*, 47, s. 939–956.

Joanna Landmesser

BADANIE AKTYWNOŚCI EKONOMICZNEJ LUDNOŚCI WIEJSKIEJ – ANALIZA Z WYKORZYSTANIEM EKONOMETRYCZNYCH MODELI HAZARDU

W pracy analizowana jest długość czasu trwania zatrudnienia wśród ludności wiejskiej. Szacując modele hazardu wyznaczamy bezpośrednie ryzyko tego, że zatrudniony przestanie wykonywać swą pracę w danym przedziale czasowym. Modele hazardu pozwalają uwzględniać takie charakterystyki badanych osób, jak płeć, wiek, poziom wykształcenia, miejsce zamieszkania, czy status zatrudnienia jako determinanty dla prawdopodobieństwa zaprzestania zatrudnienia. Wskazujemy różnice w aktywności ekonomicznej ludności wiejskiej w porównaniu z miejską oraz identyfikujemy różnice pomiędzy województwami. W analizie wykorzystujemy dane z Badania Aktywności Ekonomicznej Ludności Polski.