Survival Analysis on Employment Instability in South Korea

Hyunjin Jo[1]

GiSeung Kim[2]

-Abstract-

Recent labor market has enlarging flexible employment. Unlike past, the concept of permanent job has been diluted and it is used various employment types such as regular, temporary, service contract and part-time employment in labor market. Especially, the average of continuous service years of Korea was lowest in OECD countries. It means that voluntary or involuntary turnover rate of Korean workers is high. If it has the high flexibility in labor market, it is possible to increase the diversity of using workforce but it may cause job insecurity. Income is affected by job instability and this may cause income insecurity, which could act as a zoom factor of income inequality. Therefore we will analyze extent of employment instability in domestic labor market using a KLIPS(Korea Labor and Income Panel Study) data of 13 years, and find the relation between income and job instability through survival analysis.

[1] Ph. D candidate, Department of Economics, Pusan National University, Busan, South Korea. dora2mon@naver.com

[2] **(Corresponding Author)** Professor, Department of Economics, Pusan National University, Busan, South Korea. gsk@pusan.ac.kr

I Introduction

Employment concepts and types have been changed along with industry and socio economic situation. In Korea, workers tended to take lifelong employment for granted at early industrial era. But since 1987, labor relations made internal labor market more rigid. Since 1997 IMF crisis, restructuring and regular employment adjustment was appeared and the firms have sought cost saving and employment flexibility through the expansion of indirect or temporary employment (Choi, 2009). These socio-economic changes and the economic crisis were an opportunity for companies to operate elastically labor costs based on labor market flexibility and it has brought a lot of changes in employment way.

Regular job ensuring the retirement age is a still most popular type of employment but many people occupy irregular area in labor market such as temporary, daily or contract workers willingly or reluctantly. It is a well-known that the changes of employment type lead to job instability. However, we need to not only focus on job instability resulted from employment type change but also find out how socio-economic and human factors affect employment insecurity in the situation that labor market flexibility has dominant way.

In the meantime, some researches of job instability related with economic fluctuations such as IMF crisis (1998) or financial crisis (2008) were conducted in Korea (Cho and Keum, 2001; Lee, 2010; Kim and Kim, 2013). Cho and Keum(2001) analyzed job insecurity by gender, employment type, age, industry and occupation using the KLIPS(Korea Labor and Income Panel Survey) data and investigated the effects of age, gender and education etc. on job insecurity. They confirmed that the degree of instability looks very sharply rise after the IMF crisis. In addition, there are some studies about unstable youth employment (Lee and Kim, 2003; Woo and Yoon, 2008) or comparison of full-time job and irregular job instability (Han and Chang, 2000). Targeting local labor market, Jang and Lee (2013) conducted

regional comparative studies about precarious employment. Han and Jang (2000) argued that there need to set up new approach to job insecurity caused by the irregular employment analyzing whether to appear irregular employment temporarily by turnover etc. in the labor market or experience irregular employment continuously in certain populations groups such as women or elderly people. Jang and Lee (2013) analyzed any difference of job insecurity in regions in Korea considering that diverse characteristics of the regional labor markets.

The studies conducting the job instability seem the difference in the subject, analysis method and data. Also defining job instability is some different. The concept and measuring method of job insecurity or instability had not been suggested consistently until the early 1990s. But finally in 1997, the concept of labor market instability became established by Diebold et al. based on “How long engaged in the current employer or a job?” which is survey items on the CPS (Current Population Survey) (Cho & Keum, 2001). The definition of terms of employment insecurity and the measurement are presented variously in the earlier studies. Regular employees’ retention rate (Kim and Kim, 2013) or job retention (Cho and Keum, 2001) were used for the study and job retention rate was calculated based on years of service.[1]

This study defines the concept of job instability based on survey items in CPS to period of continuous service and measures job instability using survival analysis method. We use the period of continuous service data from KLIPS. After analyzing the probability of survival, we examine the impact of human factors such as gender, age and education level or socio-economic factors such as wages, industry and region on survival rate. We will describe the data used for survival analysis and methodology in chapter II, will present the result in

[1] Two terms are typically looks similar word but calculation method is different obviously. If you want to know more details, please refer to the two papers.

chapter III and will find out the conclusion and implication in final chapter.

II Methodology

1. Method

In many fields, researchers want to know time duration to until specific event happen to subject, and they would like to analyze the factors affecting the event. Survival analysis is one of the most widely used analytical methods in the above cases. We define retirement as event, and observe the period of continuous service for each individual. Subsequently, we will carry out an analysis of the influence factors on the event.

Therefore, in this paper, survival function refers to probability of maintaining a continuous service without retirement. Survival time variable, namely the duration of continuous service represents by T and real observed duration of continuous service represents by t. Survival function is as following.

$$S(t) = Pr(T \geq t)$$

As the contrary concept of survival function, we use cumulative distribution function which means the probability of occurring event by time t. This is as following.

$$F(t) = Pr(T \leq t) = 1 - S(t)$$

Density function (f(t)) is used to find the intensity on the case at a given point in time. It measure the risk of incidents in the total number of the entire analysis period. In contrast, hazard function (h(t)) means a risk for persons who survived to the time ‘t’. Therefore, this is the conditional probability that calculating the event observed in short time from time ‘t’. These are as following.

$$f(t) = \lim_{\Delta t \to 0} \frac{[\Pr(t \leq T < t + \Delta t)]}{\Delta t}$$

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t}$$

All of the expression described above is linked with each other. Thus, if we know one of these values, we can calculate the rest easily. We estimate the survival function for the duration of the workers on the basis of the above mathematical model.

To analyze the influence factors for the survival probability, we use the Cox-proportional hazard model. There are several reasons that the Cox regression is so important and widely applied in biomedical, engineering, economic, social, and behavioral sciences. First, prior to the Cox regression, the leading approach to multivariate survival analysis was the parametric model. But the Cox regression is a distribution-free model and does not require such information. Second, the main estimation method Cox developed is partial likelihood, which is innovative in several ways. Among other things, it allows the user to estimate the regression coefficients of the proportional hazards model without having to specify the baseline hazard function, and the estimates depend only on the ranks of the event times, not their numerical values. Third, the Cox regression is the very first model that permits the user to incorporate the time-varying covariates in survival analysis (Guo, 2010).

We find the factors affecting employment instability in accordance with wage. In addition to this, we consider human characteristics, regional properties, and industrial features.

2. Data

We used KLIPS (Korea Labor and Income Panel Study) data. This include personal data and work history for 13 years. We merged each data set into one large data set and obtained

number of 184,790 raw data. The process of erasing almost useless data is as following. First, we kept common data in personal and work history data set, and second, we deleted retrospective data because it doesn't indicate accurate date of entering or retiring a company. Third, we eliminated the data which is shown that person have own job. At last, we erased some incorrect data or unknown data, and finally we used number of 21,916 data in this analysis. It consists of retirement (event) data and censored data. The KLIPS data is consisting of wage worker and non-wage worker. We use just wage worker to analyze. The summary statistics are shown in Table 1.

Table 1 Summary Statistics

Variables		Mean	Percent (%)
Duration (days)		1408.79	-
Monthly Wage (won)		141.89	-
Age		38.69	-
Gender (no. of Man)		11,835	54.00
Education years		12.04	-
Birth Year	Before 1949	2,410	11.00
	1950-1959	3,384	15.44
	1960-1969	5,286	24.12
	1970-1979	7,108	32.43
	After 1980	3,728	17.01
Education Level	Under Middle S	4,905	22.38
	High School	9,435	43.05
	College	6,869	29.05
	Graduate School	707	3.23
Regular or non- regular[2]	Regular Job	14,258	65.17
	Non-regular Job	7,619	34.83

[2] In this paper, we define 'regular worker' as who are guaranteed retirement.

Labor Union	Existence	2,227	10.16
	Union Member	1,056	4.82

III Result

1. Life Table for working-duration data

In this part, we show several Life table for various variables we conducted. At the first, Table 2 which is below indicate average and cumulative working duration by region. In the case of Korea, average duration is almost 1408.79days and average duration of Seoul and Busan are 1355days and 1165days. They are shorter than that of Korea.

Table 2 Average and cumulative working duration by region

	Avg. duration (days)	Duration (days)					No. of Variable
		10%	25%	50%	75%	90%	
Korea	1408.79	143	303	693	1,660	3,530	21,916
Seoul	1355.42	150	304	699	1,583	4,771	5,112
Busan	1165.16	112	244	528	1,246	2,674	2,074

1) Employment Probability Estimate by regional working duration

Table 3 show survival table by regional working duration. We divided working duration into units of 6months and 1year periods. ‘Keep working’ means that periods which are workers are having own job and working. ‘Event’ means occurrence of unemployment. ‘Censored’ means cut off of observation in whole period. We shade survival rate outcomes. The working duration is longer, so survival rate is lower. By regional groups, the survival rate is lowest in Busan. For two period of 1 to 2 years and 2 to 3 years, the survival rate of Busan is as low as 0.1 compared with that of Korea.

Table 3 Survival table by working duration

		6 m	6m-1y	1-2y	2-3y	3-5y	5-10y	10y-
Korea	Keep Working	21916	18841	15014	10498	7855	4878	2049
	Event	2789	3037	3662	2034	2106	1741	1003
	Censored	173	652	808	555	834	1084	1046
	Survival rate	0.872	0.725	0.538	0.430	0.309	0.188	0.071
	Std. error	0.002	0.003	0.004	0.004	0.003	0.003	0.003
Seoul	Keep Working	5112	4423	3515	2466	1841	1110	450
	Event	631	749	890	499	539	407	233
	Censored	32	126	152	117	183	250	217
	Survival rate	0.876	0.721	0.531	0.420	0.292	0.174	0.060
	Std. error	0.005	0.006	0.007	0.007	0.007	0.006	0.005
Busan	Keep Working	2074	1702	1273	841	585	349	130
	Event	351	376	365	208	169	129	67
	Censored	10	40	63	44	65	90	63
	Survival rate	0.830	0.641	0.448	0.332	0.232	0.135	0.048
	Std. error	0.008	0.011	0.011	0.011	0.010	0.009	0.008

Table 4 show cumulative survival table by regional working duration. The analysis period for this table was set up from 1 year to 30 years. Cumulative survival rate is decreased with time. In this case, the rate is lowest in Busan through whole period.

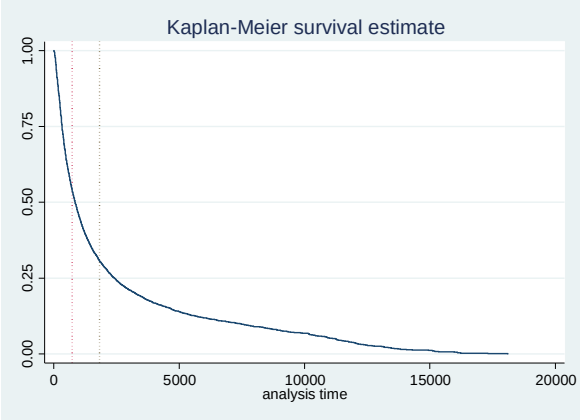
Table 4 Cumulative survival table by working duration

	1y	2y	3y	5y	10y	15y	20y	25y	30y
Korea (A)	0.719	0.539	0.433	0.316	0.209	0.172	0.157	0.148	0.142
Seoul (B)	0.716	0.532	0.423	0.298	0.191	0.152	0.138	0.129	0.124
Busan (C)	0.634	0.448	0.335	0.236	0.150	0.130	0.120	0.111	0.106
A-B	0.003	0.007	0.010	0.018	0.018	0.020	0.019	0.019	0.018
A-C	0.085	0.091	0.098	0.08	0.059	0.042	0.037	0.037	0.036
B-C	0.082	0.084	0.088	0.062	0.041	0.022	0.018	0.018	0.018

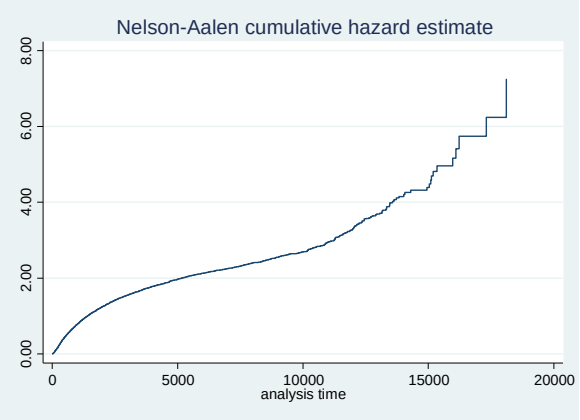
Figure 1 show unemployment probability estimates. We can recognize how rapidly working probability decrease by identifying the slope of each graph. In generally smoothed hazard estimate graph indicates the risk rate at the moment, and it shows moment rate of change from worker to unemployed in this analysis. Looking at the smoothed hazard estimate graph during 10years, the slope sharply increases until one year and declines until two years and thereafter it decreases gradually. Moment unemployed hazard rate in Busan is very high compared to that of other regions. In the case of cumulative hazard graph, the failure rate of Seoul and Busan is higher than that of Korea.

Because moment unemployed hazard rate of Busan is higher than that of other areas in more short time, we can consider that the survival rate of Busan is low as shown Table 3 and 4.

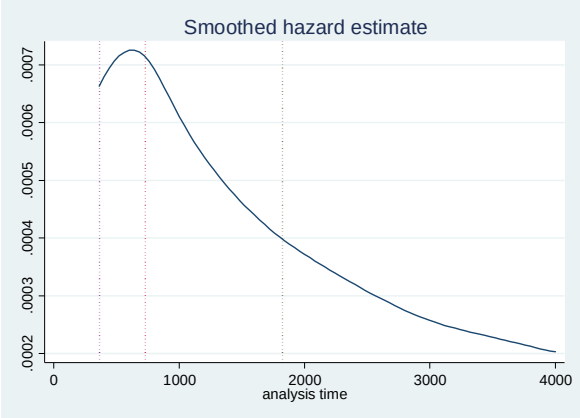
Figure 1 Unemployment probability estimates in Wage Workers



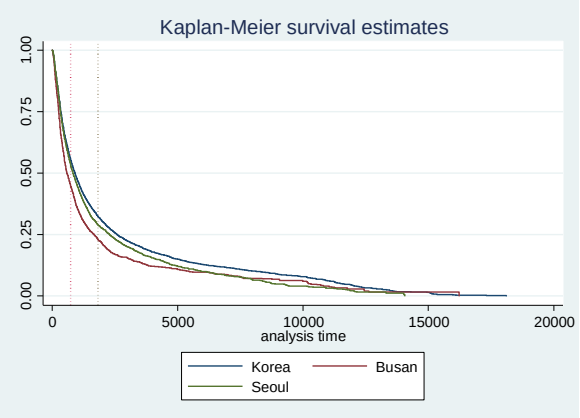
K-M Survival estimates of Korea



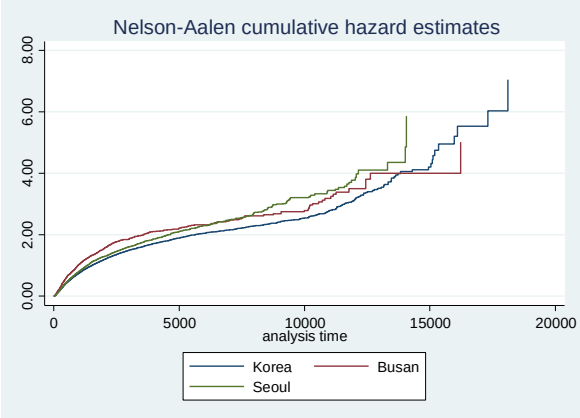
Smoothed hazard estimates of Korea



K-M Survival estimates by region

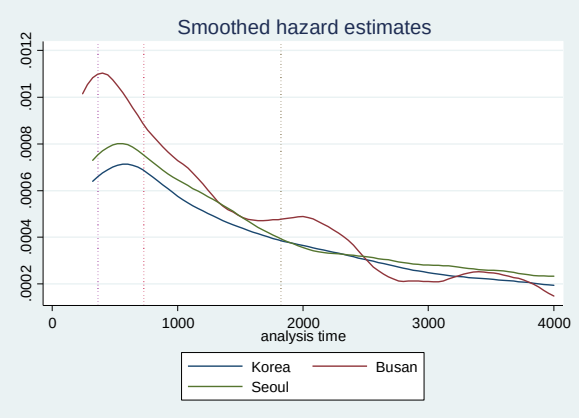


Smoothed hazard estimates by region



Nelson-Aalen Cumulative hazard estimates

by region



Smoothed hazard estimates by region

(during 10 working-years)

2) Employment Probability Estimate by working duration, education level

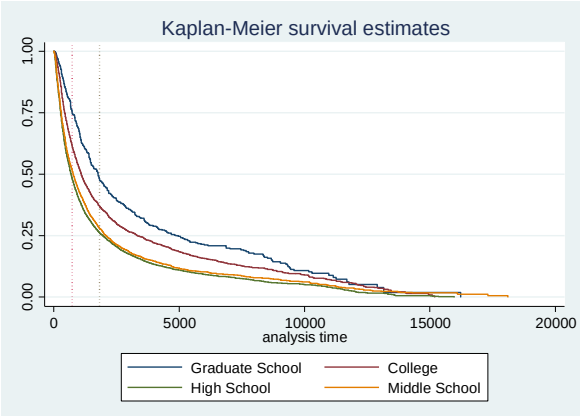
We show cumulative survival table by education level in Table 5. Graduate school level has the highest cumulative survival rate throughout the entire period and the following is College level. In Figure 2, the moment unemployed hazard rate is highest at High school level in overall education level. The peak of hazard rate appears late in College level and Graduate school level unlike those of others. And the deviation of the moment hazard in Graduate

school is small. Moment unemployed hazard rate reaches its peak between 1-2 years in Middle school, High school and College level. Thereafter they are decreased rapidly until about 5 years and declined gradually after that time. In the case of Graduate school, gradual decrease shown until 5 years has similar trajectory but the size of peak is very different. This can be considered that because the proportion of workers who have graduate school level is high in profession, they have relatively stable employment conditions.

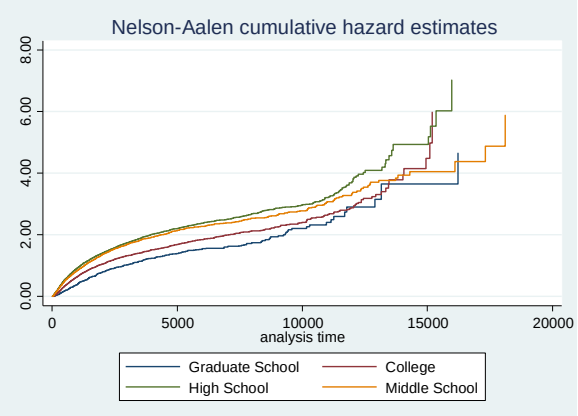
Table 5 Cumulative survival table by working duration, education level

	1y	2y	3y	5y	10y	15y	20y	25y	30y
Middle	0.694	0.516	0.410	0.290	0.184	.0149	0.138	0.129	0.122
High	0.668	0.482	0.380	0.272	0.175	0.145	0.133	0.124	0.120
College	0.788	0.614	0.504	0.380	0.262	0.217	0.199	0.189	0.183
Graduate	0.886	0.748	0.637	0.491	0.336	0.265	0.241	0.208	0.190

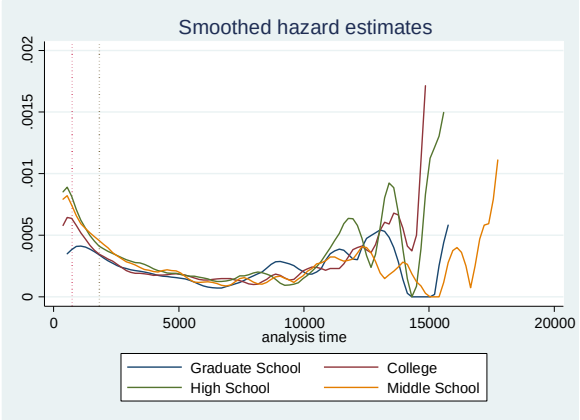
Figure 2 Unemployment probability estimates by Education level in Wage Workers



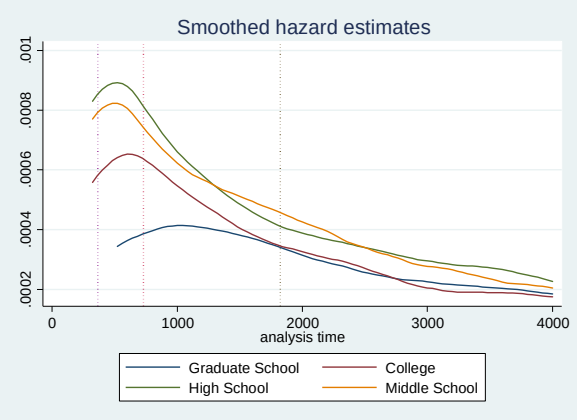
K-M survival estimates by education level



Cumulative hazard estimates by education level



Smoothed hazard estimates by education level



Smoothed hazard estimates by education level

(during 10 working-years)

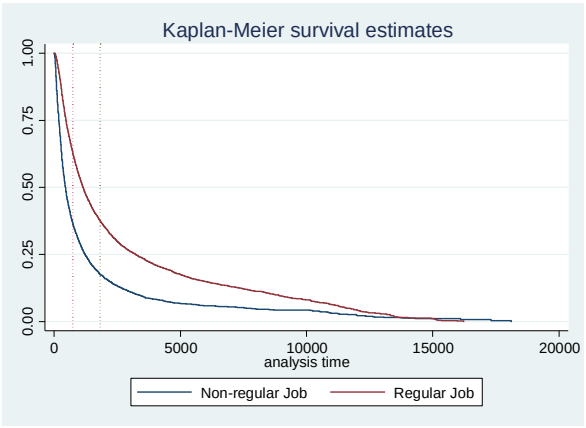
3) Employment Probability Estimate by working duration, regular job

We show cumulative survival table by regular job in Table 6. We define full-time worker as regular job and temporary worker and daily worker as non-regular job using classification of employment status of KLIPS. In Table 6, the survival rate of regular job is higher than that of non-regular worker in whole period. The gap is reduced over time. However survival rate of regular job is almost twice as high for 10 years. In smoothed hazard estimate of Figure 3, hazard rate of non-regular job reaches its peak at 1 year and the rate is declined until 2 years and thereafter decreased gradually. In the case of regular job, the peak of hazard rate is half or less of non-regular job and the rate is gradually decreased throughout whole period.

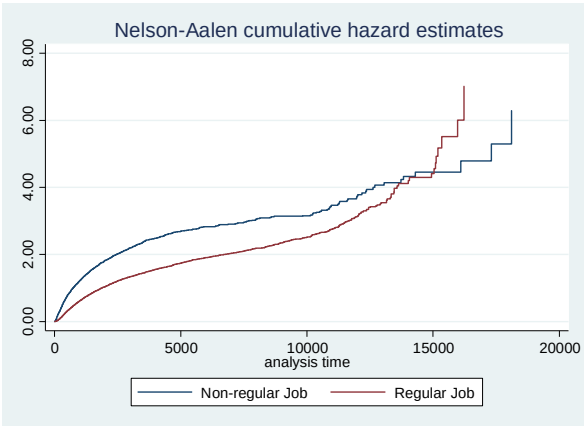
Table 6 Cumulative survival table by working duration, regular job

	1y	2y	3y	5y	10y	15y	20y	25y	30y
Regular (A)	0.808	0.631	0.516	0.381	0.251	0.201	0.182	0.168	0.161
Non-regular (B)	0.550	0.366	0.277	0.195	0.132	0.118	0.114	0.112	0.108
A-B	0.258	0.265	0.239	0.186	0.119	0.084	0.068	0.057	0.052

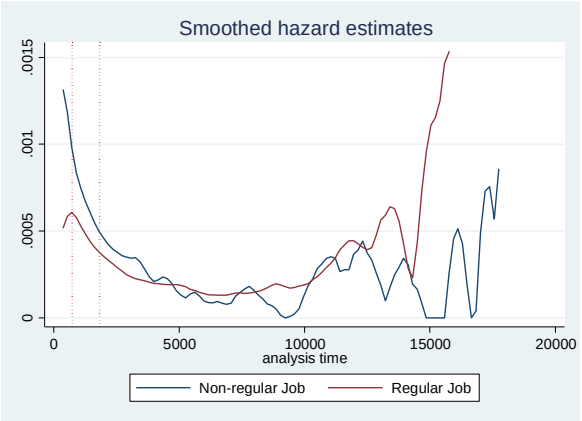
Figure 3 Unemployment probability estimates by Regular Job in Wage workers



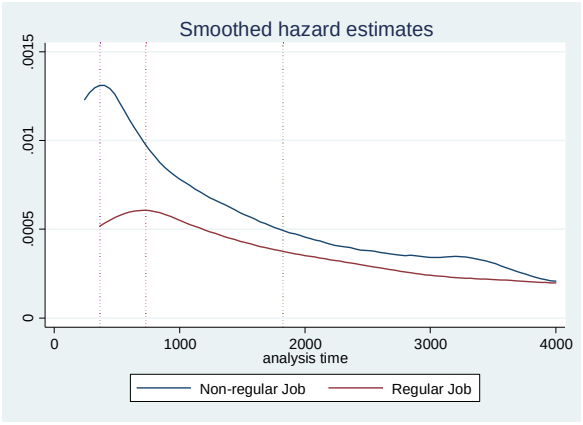
K-M survival estimates by Regular Job



Cumulative hazard estimates by Regular Job



Smoothed hazard estimates by Regular Job



Smoothed hazard estimates by Regular Job
(during 10 working-years)

2. Cox-proportional Hazard Regression

We conduct cox-proportional hazard regression by model 1 and model 2. Model 1 includes personal characteristics and model 2 includes the characteristics of labor market. Education level is included in both models. All variables shown significant at the 1% wald significance

level.

The result by birth year in model 1 is following. When we set birth year in 1970s as the reference variable, we can check that the hazard ratio is reduced in high age group. The hazard ratio of workers born before 1949 is higher than that of those born in 1950s. The hazard ratio of youth group who was born after 1980 is highest in all age groups. It is almost 40% higher than 1970s group's. This can be seen to be due to the presence of the instability of the youth labor market such as the issue of youth unemployment.

And next watching education level, we set High school level as the reference variable in education level. The hazard ratio of College level and Graduate school level is lower than that of High school level. If someone has College level or Graduate School level, his or her hazard ratio is almost 60% lower. In the case of education level variable, the hazard ratio gap depending on education level is bigger in model 2. For Middle school lever, the hazard ratio is similar with model 1. However, for College level and Graduate school level, the hazard ratio have decreased by 20%p. More highly educated workers have high probability of regular job and they are working in long tenure area such a professions and maintaining a relatively stable employment.

We define the average monthly wage of less than 1,500 thousand won to low wage. The hazard ratio of workers get low wage is more than twice that of non-low-wage workers. We consider this result that low-wage workers are comprised of women and older people who are working in irregular worker such as temporary worker or daily worker including young people who are engaged in part-time job or who cannot seek good quality job.

In the case of regular job, we set full-time worker as the regular job and reference variable. The result is as follows. The temporary worker has 1.7times hazard ratio in comparison to regular job and hazard ratio of daily worker is 1.3 times of regular job. If there is the labor

union, the hazard ratio is reduced by 60% compared with the case without labor union.

Above results suggest that the unemployment risk is determined on various factors such as age, wage, education level, employment status and labor union and there is significant difference depending on variables.

Table 7 Cox-proportional hazard regression

Variable	Model 1			Model 2		
	Hazard Ratio	Robust Std. Error	Coefficient	Hazard Ratio	Robust Std. Error	Coefficient
Gender	0.761	0.012	-0.273***			
Age				0.899	0.004	-0.107***
Age^2				1.001	0.000	0.001***
Before 1949	0.717	0.022	-0.333***			
1950-1959	0.600	0.016	-0.510***			
1960-1969	0.759	0.016	-0.276***			
After 1980	1.445	0.035	0.368***			
Middle School	1.134	0.027	0.126***	1.160	0.030	0.149***
College	0.653	0.012	-0.426***	0.855	0.016	-0.157***
Graduate School	0.598	0.025	-0.514***	0.842	0.037	-0.173***
Low-Wage[1)]				2.305	0.045	0.835***
Temp. Job				1.712	0.037	0.538***
Daily Job				1.374	0.041	0.318***
Union				0.621	0.015	-0.477***
Wald chi^2	1754.08			7768.72		
Log pseudolikelihood	-151893.77			-148273.76		

Note:
1. Low-wage means that average monthly wage is less than 1,500 thousand won.
2. Mark '***' indicates 0.01 significant level.

IV Conclusion

This analysis is conducted to confirm instability of labor market and influence factors by survival analysis using 13-years (from 1998 to 2010) data of KLIPS(Korea Labor and Income Panel Study). By region, we analyzed data of Korea, Seoul and Busan. And by detailed variables, we used education level, regular job and householder variables. Also we checked determinant factors affecting survival rate using cox-proportional hazard regression.

The results described briefly as follows. First, average working duration of wage worker in Korea is 1408.79 days(almost 3.86 years) and those of Seoul and Busan are 1355.42 days (almost 3.71 years), 1165.16 days(3.19 years) respectively. The working duration of workers in Busan is shorter than that of Korea and Seoul and survival rate is low too. We can explain that because moment unemployed hazard rate of Busan is much higher than that of other areas in more short time. Second, full-time workers classified regular job has higher survival rate in the long term than temporary workers or daily workers classified non-regular job. And moment unemployed rate of regular job is very low and gradual so we can consider that they are maintain relatively stable employment. Third, more educated workers has high survival rate and they are not affected largely in the short term. So we can consider that they have relatively stable employment. Finally, Low-wage has a very large impact on hazard ratio. The determinants factor affecting employment instability of wage worker is varied according to the personal characteristics or labor market characteristics. The degree of each influence factors is appeared differently too.

What can we learn from the results of this study are as follows. The impact of structure in the labor market such as regular and non-regular job is comparatively big and education level has significant impact. Of course, this suggest the well-known phenomenon that reliability and stability of regular job or profession of highly educated people. However our result is very

important because these problems have been formed in the labor market during the period of more than 10 years. If we do not change the structure of such a labor market, this difference can be even greater and can be more intensified. The hazard ratio of low-wage worker is

Depending on the expansion of the modern labor market flexibility, increase of non-regular workers can be unavoidable. But job seekers including new entrants are still prefer regular job to non-regular job due to unstable employment of non-regular job. For that reason, there can be permanent problem which is repeated vicious circle that lots of people are driven into decent job and low-quality job is vacant. Therefore, stability of employment of non-regular worker is paramount. In addition, it is needed wage increase and betterment of labor condition to ensure more stable employment for non-regular job, low-educated workers and low-wage workers. We suggest that support for the structural improvement of the labor market and efforts are needed on an ongoing basis.

Reference

Choi, Young-Ki. "Economic Crisis and Korean Social Model." *Korean Journal of Industrial Relations, Vol. 19, No. 3.* 2009. 119-149.

Cho, Joonmo and Jaeho Keum. “Job Instability in the Korean Labor market: Comparison before and after the IMF Economic Crisis.” *Korean Journal of Labour Economics,* Vol. 24, No. 1. 2001. 35-66.

Cleves, Mario, Roberto Gutierrez, Willian Gould, and Yulia Marchenko. 『An Introduction to Survival Analysis Using Stata.』 Stata Press, 2008.

Coetzee, Celeste. "A practical guide to survival data analysis for panel surveys using stata." centre for social science research, College of Cape Town.

Cox, D. R. "Regression Models and Life-Tbles." *Journal of the Royal Statistical Society. Series B (Methodological), Vol. 34, No. 2.* . Royal Statistical Society, 1972. 187-220.

Guo, Shenyang. *Survival Analysis.* Oxford College Press, 2010.

Han, Joon and Ji Yeun Chang. “Work history and Life-Course by Regular/Irregular Switching.” *Korean Journal of Labour Economics*, Vol. 23, 2000. 33-53

Heckman, J. and B. Singer. "The Identifiable of the Proportional Hazard Model." *The review of Economic Studies*, Vol. 51, No. 2. 1984. 231-241

Hosmer, David, Stanley Lemeshow, and Susanne May. 『Applied Survival Analysis.』 John Wiley & Sons, Inc., 2008.

Jeong, In Soo, Byeong Yoo, Jeon and Sang Hun Lim. “Local Labor Market Research.” Korea Labor Institute, 2003.

Jang, Yeon-Joo and Min-Hwan Lee. “Estimating the condition and determinants of wage worker’s job instability by regions.” *Journal of Regional Study*, Vol. 21, No. 1, 2013. 121-145

Kim, Sunghwan, and Jong-Woon Kim. "Effects of Firms' Financial Stability on Job Creation and Stability in Korea." *Journal of Business Research*, Vol. 28, No. 4. 2013. 271-298.

Lee, Byoungki. *The Survival of New Firms in Korean Manufacturing: Start-up Conditions and Post-Entry Evolution*. KERI, 2003.

Lee, Eul-Teo. "Historical Review on the Relationship Between Political Enonomic Environment and Employment Relations System." *The Review of Business History, Vol. 25, No. 4.* 2010. 207-242.

Lee, Kyu-Yong, Yong-Hyun Kim. “The determinants of Youth Labor Market Performance.” *Labor Policy Research*, Vol. 3, No. 2. 2003. 69-94

Macran, Susan Heather Joshi and Shirley Dex. "Employment after Childbearing: A survival analysis." *Work employment and Society*, Vol. 10, No. 2. 1996. 273-296.

Park, Jaebin. *Survival Analysis, theory and practice.* Shinkwang Pub., 2006.

Park, Jin Kyung, Kwang Ho Oh and Min Soo Kim. “Survival analysis on the business types of small business using Cox’s proportional hazard regression model.” *Journal of the Korean Data and Information Science Society*, Vol. 23, No. 2. 2012. 257-269.

Shumway, Tyler. "Forecasting Bankruptcy More Accurately: A Simple Hazard Model." *Journal of Business*, Vol. 74, No. 1, 2001. 101-124.

Woo, Hae-Bong, Yoon, In-Jin. "An Analysis of First Job Attainment among Young Koreans, Utilizing the Competing-Risks Survival Analysis Model." *Studies on Korean Youth*, Vol. 19, No. 2. 2008. 225-250.