Japanese Wage Curve: A Pseudo Panel Study

Shigeki Kano †

May 2003

†Correspondence: Shigeki Kano, Doctoral Program in Policy and Planning Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-8573, Japan (e-mail address; skano@sk.tsukuba.ac.jp). The author is grateful to Professor Kanemi Ban for kindly providing the data used in this paper. Professor Makoto Ohta is also acknowledged for his helpful comments.

Abstract

We provide the first empirical evidence of the Japanese wage curve in a comparable way with previous US and European cases, using the pseudo panel of totally 5091 cohorts for 1984, 1988, and 1994. The Japanese wage elasticity with respect to regional unemployment is statistically significantly negative. This can be explained by a peculiar wage setting process in Japan.

JEL classification: C23; J31; J60

Keywords: Wage curve; Regional unemployment; Pseudo panel data

1 Introduction

In their seminal book, Blanchflower and Oswald (1994) advocate a new economic stylized fact; a negative relationship between individual wages and regional unemployment, termed the *wage curve*. They test this negative relationship for 12 countries employing their large microeconomic panels in a unified manner, and find that the estimated individual wage elasticity with respect to regional unemployment is approximately -0.1 for almost all the countries.[1]

This finding has been recognized to be of particular importance, since unemployment and wages are positively associated by traditional labor market models which have an upward sloping labor supply curve and downward sloping labor demand curve. Blanchflower and Oswald (1994) point out that such a negative wage-unemployment relationship can be implied by union-firm bargaining or efficiency wage models. More recently, Sato (2000) shows that equilibrium unemployment models are also consistent with the wage curve. Therefore, the existence of the wage curve indicates non-competitive features of labor markets.

Further, this observed negative relationship suggests that the status (tightness) of the labor market is one of the important determinants of employed workers' ernings. Traditional studies of estimating earning equations based on personal characteristics, e.g., age, gender, and educational levels, seems to have ignored this viewpoint (see, for example, Heckman and Polachek, 1974).

The purpose of this paper is to provide the first statistical evidences of the Japanese wage curve in a comparable way with previous studies, i.e., employing a large microeconometric panel data.[2] It is well-known that the Japanese labor market has remarkably different institutional

[1]The countries covered by their study are USA, Britain, Canada, Korea, Austria, Italy, Netherlands, Switzerland, Norway, South Ireland, Australia, and Germany.

[2]Blanchflower and Oswald (1994) refer to Montgomery (1993) as an example of the Japanese wage curve. However, his work cannot be compared directly with other existing microeconometric studies in that it is based a on more aggregated, prefecture-level data.

characteristics from previously studied countries. It is therefore worthwhile testing the robustness and applicability of Blanchflower and Oswald's (1994) finding by using the Japanese data.

In Japan, however, individual data sets are rarely available due to a strict privacy protection law on the use of survey data collected by the government's survey. So we utilize a three year pseudo panel recently created by Ban and Takagi (1999). This data contains totally more than 5000 of appropriately defined cohorts for 1984, 1989, and 1994. Verbeek and Nijman (1992) show that within estimates based on a pseudo panel can be viewed reasonably as these based on a genuine panel when the number of individuals in each cohort is sufficiently large, and our data meets this condition. Hence, the result of our pseudo panel study estimating the wage curve can be compared with previous individual panel results.

The rest of this paper is organized as follows. Section 2 specifies the wage curve. Section 3 describes the data used in our analysis, and shows estimation results. Section 4 concludes the paper.

2 Specification

Following Blanchflower and Oswald (1994), we specify the wage curve as a log-linear form. For $c = 1,2,...,C$, $r = 1,2,...,R$, and $t = 1,2,...,T$,

$$\ln(W_{c,r,t}) = \mu_c + \lambda_t + D_r + \beta \ln(U_{r,t}) + \gamma Z_{c,t} + \varepsilon_{c,r,t}, \tag{1}$$

where $W_{c,r,t}$ denotes a wage payed to cohort c who is a resident of region r in period t, while $U_{r,t}$ the unemployment rate of a region where cohort c resides. D_r denotes a dummy variable indicating region r. Intercepts μ_c and λ_t denote cohort and time effects controlling unobservable components. Previous studies often do not use individual specific effects, but we do use cohort effects for a reason explained below. $Z_{c,t}$ denotes other control variables which are not of interest in estimating the wage curve. It is assumed that error term $\varepsilon_{c,r,t} \sim \text{iid}(0, \sigma^2)$, but this can be relaxed. Parameter β denotes the wage elasticity with respect to regional unemployment, and $\beta < 0$ is the hypothesis to be tested in this paper.

We include only two variables in $Z_{c,t}$, the logarithm of ages and its square, due to data unavailability. This might be seemingly problematic because previous studies use occasionally more variables as controls, e.g., gender, race, educational levels, and occupational categories. Instead we include the fixed cohort effects μ_c in regression equation (1) to presumably control such time-invariant characteristics, utilizing the panel structure of our data. Also, the time effects λ_t is used for controlling common effects of nominal and real macroeconomic factors on each cohort wage, such as inflation and technical changes in production process.

Two points should be noted prior to estimation. First, the data used in this paper is a pseudo panel which consists of sample means of individuals in each cohort. Deaton (1985) shows that within estimates based on such sample cohort mean variables have bias due to their measurement errors. However, Verbeek and Nijman's (1992) Monte Carlo study reveals that this bias can be negligible, when the number of individuals in each cohort, say N_C, is sufficiently large (about 100 or more). Notice that, as pointed out by Verbeek and Nijman (1992), increasing N_C means reducing the number of available cohorts, i.e., the sample size $C \times T$. This gives rise to a trade-off between reduction of possible bias in estimates and the degree of freedom. So we estimate equation (1), using smaller N_C but larger sample size and larger N_C but smaller sample size.[3]

[3]We do not employ the error adjustments method of Deaton (1985), since cohort covariances are not available.

Next, notice that one of our explanatory variables (regional unemployment rates) is more aggregated than the dependent variable (cohort wages), and it constitutes groups in the dependent variable. Moulton (1990) shows that, in this situation, the covariance matrix should be estimated by the following formula;

$$\hat{\sigma}(X'X)^{-1}\left[1+(m-1)\hat{\rho}\right], \tag{2}$$

where X denotes the matrix of explanatory variables and $\hat{\rho} \in [0,1]$ denotes the estimate of the correlation coefficient of within group error term. Parameter m denotes the ratio of the sample size to the number of groups, which is given by $(C \times T)/(R \times T) > 1$ in our case. Therefore, conventional covariance matrix estimator $\hat{\sigma}(X'X)^{-1}$ has an obvious downward bias.

In order to deal with this problem, Blanchflower and Oswald (1994) and subsequent studies regress the regional averages of individual wages on regional unemployment rates to make m be equal to unity, which is called the "cell mean regression". We follow their somewhat crude strategy. In addition, we estimate the upper bound of the covariance matrix by assuming $\rho = 1$ and estimating $\hat{\sigma} = \hat{\varepsilon}'\hat{\varepsilon}/(C \times T - K)$, where K denotes the number of explanatory variables. This gives the lower bound t-values in the presence of grouped dependent variable.[4][5]

[4] We tried to estimate ρ as well as σ^2, employing the maximum likelihood method for error components models described by Searle (1971, page 462-463). However, our GAUSS code of Newton-Rhapson or BHHH numerical maximization algorithm failed to converge after thousands of iterations. This computational failure seems due to our extremely large sample size, more than 5000.

[5] Scott and Holt (1982) show that, generally, the estimator of σ depend on that of ρ, but conventional estimator of σ, i.e., $\hat{\varepsilon}'\hat{\varepsilon}/(C \times T - K)$ is consistent even in this situation.

3 Data and Estimation Results

The pseudo panel data used in this study is given by Ban and Takagi (1999). This data is based on *National Consumer Survey (Zenkoku Shôhishya Zittai Chôsa)* for 1984, 1989, and 1994, which is conducted by Statistic Bureau of the Ministry of Public Management, Home Affairs, Posts and Telecommunications, Japan. The cohorts in this pseudo panel are constructed as follows. First, the national observation is divided into 47 prefectures. Next, each prefecture is further divided into ages by a year old. Ban and Takagi (1999) provide sample cohort means and sample cohort variances of variables (mainly of consumption and saving categories) for surveyed worker establishments.

Originally, Ban and Takagi's pseudo panel contains 1804 cohorts for 1984, 1892 for 1989, and 1998 for 1994. However, in order to avoid problems concerning unbalanced panels, we drop the cohorts with missing years from the sample we investigate. Consequently the total number of cohorts used, i.e., the sample size of data in this study, is reduced to 1697 (cohorts) $\times$ 3 (years) = 5091 (See Table 1). We use net cohort annual income as $W_{c,r,t}$, which is deflated by the regional price index of Statistic Bureau. However, it is found that similar estimates are obtained even though regionally non-deflated income is used.

The data of the unemployment rates of 47 prefectures is drawn from *Population Census (Kokusei Chôsa)* for 1985, 1990, and 1995, by Statistic Bureau. This census is the only source of the regional labor force population based on a large scale survey. Notice that they lead the aforementioned pseudo panel by a year. However this seems not to cast serious problem, because Blanchflower and Oswald (1994) show that the effect of a year lag of unemployment on the wage is approximately as same as that of contemporary unemployment.

Table 2 shows our estimation results of equation (1). Column (A) of this table reports the within estimate of wage elasticity with respect to unemployment $\hat{\beta}$, without regional dummies. On the other hand, column (B) does the within estimate with regional dummies. For both cases, the estimates are statistically significant and exhibit an expected sign, i.e., negative. Comparing the estimates in (A) and (B) brings out that the inclusion of regional dummies has not apparent impact on the value of estimates. Blanchflower and Oswald (1994, Table 8.8) show that for almost all of 12 countries they studied, $\hat{\beta}$ is approximately -1. So, our estimate of $\hat{\beta} \approx -0.18$ is remarkably larger in absolute value than other countries. This suggests that Japanese wage is more sensitive to its outside labor market status.

This empirical wage curve relationship can be explained well by a peculiar, non-competitive wage setting custom in Japan, called *Syuntô* (Spring Labor Offensive); annual union-firm wage negotiations. In *Syuntô* process, labor unions propose their wage offers, comparing their collective preference on raising baseline wages with on protecting employment of union members. The unions often reveal their priority on the latter purpose when prevailing unemployment is severe, and consequently they give up the former purpose. Thus, this Japanese wage setting custom might be at the bottom of the negative wage elasticity with respect to unemployment.

In order to eliminate possible measurement errors in our pseudo panel, we estimate equation (1) using the data whose average cohort size is increased by roughening the age division from a year to five year. The result is shown in column (C) of Table 2. However, the estimate of (C) exhibits little difference from (B) which has a smaller cohort size than (C). So the measurement error bias of Deaton (1985) might be a negligible level in our case. Notice that, as expected in the previous section, t-values are substantially reduced by loosing the degree of freedom for increasing the average cohort size.

The result of cell mean regression is given by column (D) of Table 2. Even using more aggregated wages as dependent variable, negative relationship between wages and unemployment is still significant. Further, for column (A), (B), and (C), estimates are significant even evaluated by the aforementioned lower bound t-values ($\tilde{t}$ in the table). So our results seem robust against Moulton's (1990) problem of the downward bias in the covariance matrix estimator.

Table 3 shows within estimates of three different age groups, i.e., less than 35 (the young), from 35 to 50 (the middle), and more than 50 (the old). Their estimates are significantly negative, but apparent heterogeneous wage elasticities among these groups are found. Wages of the young age are more elastic with respect to regional unemployment, i.e., more sensitive to labor market status. In addition, when evaluated by the lower bound t-values, the estimates of the middle and the old are insignificant. This result may reflect the fact that the younger workers are not fully protected by the Japanese lifetime employment system, and that there exist stronger determinants of the middle and the old workers' wage levels under this system, e.g., how long they belong to their employers with honesty.

4 Conclusion

This paper estimates the Japanese wage curve using pseudo panel of totally 5091 cohorts in 1984, 1989, and 1994. It is found that there exist a statistically significant wage curve, i.e., a negative relationship between individual wages and regional unemployment. This can be explained by a peculiar, non-competitive wage setting process in Japan, *Syuntô*.

References

Ban, K., Takagi, S., 1999. Experimental construction of pseudo panel data (gizi panel data no sakusei zikken, in japanese). Report on the Financial Support for Scientific Researches.

Blanchflower, D. G., Oswald, A. J., 1994. The Wage Curve. The MIT Press, Cambridge, MA.

Deaton, A., 1985. Panel data from time series of cross-section. Journal of Econometrics 30, 109–126.

Heckman, J. J., Polachek, S., 1974. Empirical evidence on functional form of the earning-schooling relationship. Journal of the American Statistical Association 69, 350–354.

Montgomery, E. B., 1993. Pattern in regional labor market adjustment: The ”united states” vs. ”japan”. NBER working paper 4414.

Moulton, B. R., 1990. An illustration of a pitfall in estimating the effects of aggregate variables on micro unit. Review of Economics and Statistics 72, 334–38.

Sato, Y., 2000. Search theory and the wage curve. Economics Letters 66, 93–98.

Scott, A., Holt, D., 1982. The effect of two-stage sampling on ordinary least squares methods. Journal of the American Statistical Association 77, 848–854.

Searle, S. R., 1971. Linear Models. Wiley, NY.

Verbeek, M., Nijman, T. E., 1992. Can cohort data be treated as genuine panel data? Empirical Economics 17, 9 – 23.

Table 1: Sample Size of Ban and Takagi's Pseudo Panel

Year	Surveyed establishments	Cohorts	Actually used
1984	31097	1804	1697
1989	34532	1892	1697
1994	36488	1998	1697

Source: Ban and Takagi (1999).

Table 2: Within Estimates of Wage Curve

	(A)	(B)	(C)	(D)
$\ln(U_{c,t})$	-0.180	-0.174	-0.184	-0.273
(t-value)	(-18.581)	(-13.577)	(-8.661)	(-5.843)
($\tilde{t}$-value)	(-3.281)	(-2.356)	(-3.647)	-
Regional dummies	No	Yes	Yes	-
Adjusted R^2	0.460	0.491	0.588	0.577
Sample size	5091	5091	1134	141
Ave. cohort size	19.47	19.47	88.98	862.00

Note: Cohort and time effects are included in all estimations. Only estimated coefficients of $\ln(U_{c,t})$ are reported. t-values are in parentheses, based on the heteroskedasticity-consistent covariance estimator. $\tilde{t}$ denotes the lower bound t-values calculated by Moulton's (1990) formula with $\rho = 1$.

Table 3: Within Estimates of Wage Curve by Age Group

	Age < 35	$35 \leq$ Age < 50	$50 \leq$ Age
$\ln(U_{c,t})$	-0.175	-0.047	-0.106
(t-value)	(-6.582)	(-2.855)	(-2.929)
($\tilde{t}$-value)	(-2.425)	(-0.835)	(-0.875)
Adjusted R^2	0.578	0.638	0.472
Sample size	1212	2112	1767
Ave. cohort size	15.45	25.32	15.24

Note: Cohort effects, time effects, and regional dummies are included in all estimations. Only estimated coefficients of $\ln(U_{c,t})$ are reported. t-values are in parentheses, based on the heteroskedasticity-consistent covariance estimator. $\tilde{t}$ denotes the lower bound t-values calculated by Moulton's (1990) formula with $\rho = 1$.