No. 1001

Long-Run Matching Relationship in the Japanese Labor Market: A Panel Cointegration Approach

by

Shigeki Kano and Makoto Ohta

July 2002

Long-Run Matching Relationship in the Japanese Labor Market: A Panel Cointegration Approach

Shigeki Kano
Doctoral Program in Policy and Planning Sciences
University of Tsukuba†
and
Makoto Ohta
Institute of Policy and Planning Sciences
University of Tsukuba

July 2002

†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 data set and GAUSS programming code used in this paper are available upon request.

Abstract

This paper investigates the long-run relationship among new hiring, unemployment (job seekers), and unfilled vacancies in Japan, using an annual panel data on 48 prefectures for 1972-1999. We find that in the panel data framework, these three variables are $I(1)$ processes, and are cointegrated. Further, we estimate the panel cointegration equation derived from a Cobb-Douglas matching function by the heterogeneous fully modified OLS and heterogeneous dynamic OLS. The estimation results show that conventional within estimates could have non-negligible biases.

JEL Classification Number: C23, E24, J41, J60.

1 Introduction

In the theory of equilibrium unemployment, the matching function relates unemployment and unfilled vacancies positively with new hiring. Since a prominent study by Blanchard and Diamond (1989), dozens of authors have estimated the matching functions of various countries, and they provide statistical evidences on the existence of stable matching relationship[1].

The previous studies estimated the matching function usually by using time series data, but they did not thoroughly investigate the stationarities of new hiring, unemployment, and vacancies. The stationarity is important, since it is well known that if the time series data in question is non-stationary, then conventional techniques of estimations and tests are no longer applicable.

Unemployment, one of the explanatory variables in the matching function, is often characterized as a unit root process in existing studies on macroeconomic time series analysis[2]. The term "hysteresis" has been frequently put on unemployment series for expressing its high persistency. Therefore estimating the matching function involves the possibility of non-stationarity problem.

Gross (1997) might be the first and the only accessible study which sheds light

[1]See a concise survey by Petrongolo and Pissarides (2001) for the list of previous empirical studies estimating the matching function. However, it does not include any studies on Asian countries such as Japan. One of our contributions is to add the Japanese case to the list.

[2]The time series behavior of unemployment still generates vital debates as well as other macroeconomic variables. See, for example, Papell, Murray, and Ghiblawi (2000) for the recent argument.

on the non-stationarity problem in this empirical field. He applied the so-called Johansen procedure to the German time series, and found that new hiring, unemployment, and vacancies are $I(1)$ process and cointegrated[3]. However, despite recent increasing use of the panel data sets in estimating the matching function, the non-stationarity has not yet been examined in such a panel data framework.

Since mid 1990's, econometric theories handling non-stationary panels, that is, panel unit root tests and panel cointegration analysis, have grown rapidly. They enjoy various advantages of conventional panel data analysis over the one with pure time series dimension. Especially, adding the cross sectional dimension to the data is considered to resolve the low power problems of conventional unit root tests[4].

Exploiting these new and fruitful developments in empirical tools of non-stationary panels, this paper investigates the long-run relationship among new hiring, unemployment, and vacancies with Japanese annual prefectural panel data. We find that new hiring, unemployment (actually job seekers) and vacancies are $I(1)$ processes and cointegrated. It implies the existence of long-run matching relationship among these variables assumed by the equilibrium unemployment theory. Further, the panel cointegration equation derived from the matching function

[3]Note that his estimated cointegration vectors are far from what is considered by the equilibrium unemployment theory. He interprets these estimates as the upper and lower bounds of elasiticites of new hiring with respect to unemployment and vacancies during the sample period.

[4]See Im, Pesaran, and Shin (1997) for this explanation.

is estimated by heterogeneous fully-modified OLS and heterogeneous dynamic OLS, which we will explain briefly in section 2. It is shown that conventional within estimates are seriously biased.

The remainder of this paper is organized as follows. Section 2 introduces our estimation model and provides brief review of econometric methods used in the analysis. Section 3 shows our results of panel unit root tests, panel cointegration tests, and panel cointegration estimations. Section 4 concludes the analysis.

2 The estimation model and method

The matching function relates unemployment and unfilled vacancies positively with new hiring at a given level of the “matching technology” in a labor market. One can see how the matching function works in the equilibrium unemployment theory in, for example, Pissarides (2000).

Let $h_{i,t}$ denote the log of new hiring during period t in the i-th prefecture, and $u_{i,t-1}$ and $v_{i,t-1}$ denote the log of unemployment and vacancies at the beginning of t in the i-th prefecture, respectively. Assume that these variables are $I(1)$. Then, the following panel cointegration system is of our interest; for $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$,

$$y_{i,t} = x'_{i,t}\beta + \mu_i + \lambda_t + \varepsilon_{i,t} \tag{1}$$

$$x_{i,t} = x_{i,t-1} + \omega_{i,t}, \tag{2}$$

where $y_{i,t} = h_{i,t}$ and $x'_{i,t} = (u_{i,t-1}, v_{i,t-1})$. μ_i and λ_t in equation (1) denote individual and time effects, respectively.

$\varepsilon_{i,t}$ and $\omega_{i,t}$ in each equation denote stationary error terms with zero means. Note that the statonarity of $\varepsilon_{i,t}$ is equivalent to y_t being cointegrated with x_t by the definition of cointegration[5]. Further $\varepsilon_{i,t}$ and $\omega_{i,t}$ are assumed to have a "long-run covariance matrix", $\Omega = \sum_{s=-\infty}^{\infty} E(\epsilon_{i,s}\epsilon'_{i,0})$, where $\epsilon_{i,t} = (\varepsilon_{i,t}, \omega_{i,t})'$. Here Ω is assumed to be homogeneous among individual units, but it will be relaxed later.

The expression (1) comes from the matching function specified as a log-linear Cobb-Douglas form, which is preferably employed by previous studies. So β could be interpreted as the vector of hiring elasticities with respect to unemployment, β_u, and to vacancies, β_v, and $\beta > 0$ is a natural and testable assumption. On the other hand, (2) states that the explanatory variables have unit roots.

As in the case of pure time series models, recently Kao (1999) and Phillips and Moon (1999) show that the conventional within estimate of β is biased due to underlying endogeneity and serial correlation in the above system. In order to remove this bias, Phillips and Moon (1999) and Kao and Chiang (2000) develop alternative estimators, i.e., fully-modified OLS (FM) and dynamic OLS (DOLS)[6].

Basically, FM and DOLS are OLS with some sorts of bias correction. In the

[5]Hereafter, z_t denotes the pooled expression of variable $z_{i,t}$, i.e., $z_t = (z_{1,t}, z_{2,t}, ..., z_{N,t})'$.

[6]Kao and Chiang (2000) provides detailed and comprehensive arguments on these two estimators, including their limiting distributions, estimation procedures, and small sample performances, *etc*. Notice that in the literature of non-stationary panel data analysis, the within estimator is commonly referred to as OLS. This paper follows this tradition.

FM estimation, based on estimated long-run covariance matrix, a data transformation is done on the dependent variable y_t before running OLS. On the other hand, in the DOLS estimation, the following extra terms are added to original cointegration equation (1) so that bias is corrected, which consist of lags and leads of the first order differences of explanatory variables:

$$\sum_{j=-p_1}^{p_2} c_j \Delta x_{i,j}$$

, where c_j is a nuisans parameter to be estimated.

The assumption of homogeneous long-run covariance matrix among individual units might seem to be implausible. Putting cross sectional dimension to the regression occasionally gives rise to heterogeneity in the error term. Kao and Chiang (2000) develops the extended versions of FM and DOLS which allow the heterogeneity of the long-run covariance matrix: They are called heterogeneous FM and heterogeneous DOLS, respectively. We employ these techniques, and so hereafter the long-run covariance matrix Ω is not common to all the prefectures but depends on the individual prefecture i.

3 Estimation results

3.1 Data description

The data used in our analysis is drawn from "Referentials and Placements by Prefecture, Employment Referral Statistics" in *Year Book of Labour Statistics*, issued by The Ministry of Labour, Japan[7]. Our variables correspond to the published data as follows; h_t = log(*Placements*), u_t = log(*Active Applicants*), and v_t = log(*Active Openings*).

The sample period is 1972 to 1999, but one period lag is used for explanatory variables. So the number of time periods is $T = 27$. The number of individual units, namely of prefectures in Japan, is $N = 47$. For the details of the data set used, see Kano and Ohta (2002).

3.2 Panel unit root test

As a pre-test for the cointegration analysis, we first investigate non-stationarity of h_t, u_t, and v_t, employing the t-bar test proposed by Im, Pesaran, and Shin (1997, hereafter IPS). The IPS t-bar is designed to test H_0: all individual units have unit roots against H_1: some individual unit has not unit roots. Formally,

$$H_0 : \alpha_i = 0 \ \ \forall i, \qquad H_1 : \exists\, i \ \ \textit{such that}\ \alpha_i < 0,$$

where α_i is the coefficient of the ADF regression for each individual unit, i.e., for a time serise y_t,

$$\Delta y_{i,t} = \mu_i + \alpha_i y_{i,t-1} + \sum_{k=1}^{p} \phi_i \Delta y_{i,t-k} + \gamma_i t + \varepsilon_{i,t}, \qquad t = 1, 2, ...T,$$

where γ_i could be zero or not. In our case h_t, u_t, or v_t are assigned to y_t. Im, Pesaran, and Shin (1997) shows that the IPS t-bar consists of the average of Aug-

[7]Currently, Ministry of Health, Labour, and Welfare.

Table 1: IPS t-bar tests

without trend			with trend		
h_t	u_t	v_t	h_t	u_t	v_t
2.135	3.177	−0.292	2.740	0.767	−0.217
(0.984)	(0.999)	(0.385)	(0.997)	(0.777)	(0.414)

Note: The p-values are in parentheses. Lag length = 8 is determined by Shwert (1989)'s criterion. Note that it distributes as $N(0, 1)$.

mented Dickey-Fuller t-values for each individuals, and after appropriate normalization, IPS t-bar distributes as $N(0, 1)$ under some assumptions.

At present, there are no procedures to determine the lag length of the IPS t-bar test, p. Schwert (1989)'s criterion, which is originally proposed for pure time series analysis and depends only on T, suggests p should be 8 for our time period[8]. Invoking this criterion, we set $p = 8$ and run IPS t-bar tests with and without a trend.

The results are listed in Table 1. The table shows that the presence of unit root cannot be rejected at 5% significance level. Thus we conclude that h_t, u_t, and v_t are $I(1)$ variables, which means that the panel cointegration analysis is necessary for our data.

[8]Schwert criteria is defined as $p = p(T) = \text{Int}\left[12 \times (T/100)^{1/4}\right]$, where $\text{Int}[z]$ means the integer part of z. We also implement Ng and Perron (1995)'s data dependent lag length selection procedure for each individual unit, and find p's vary across individuals. However, this procedure cannot be applicable to IPS test, because it require p's to be the same for all individuals.

Table 2: Panel cointegration tests

DF_ρ	DF_t	DF_ρ^*	DF_t^*
−4.15	−3.728	−12.209	−7.151
(0.000)	(0.000)	(0.000)	(0.000)

Note: p-values are in parentheses. For DF_ρ^* and DF_t^*, the lag length of the Bartlett kernel in estimating Ω is 4. Note that these test statistics all distribute as $N(0,1)$.

3.3 Panel cointegration analysis

We implement Kao (1999)'s four types of residual-based panel cointegration tests, DF_ρ, DF_t, DF_ρ^*, and DF_t^*, all of which set no cointegration as a null hypothesis. These test statistics are based on Dickey-Fuller tests on OLS residuals of cointegration equation (1), and after some appropriate adjustments, all distribute as $N(0,1)$. Note that DF_ρ^* and DF_t^* need the estimate of the long-run covariance matrix, $\hat{\Omega}$[9].

The results are reported in Table 2. The null hypothesis of no cointegration is rejected by all the tests. Therefore, it turns out that there exists a long-run relationship among new hiring, unemployment, and vacancies, implied by the matching function.

Table 3 presents our estimation results of the cointegration equation (1). FM and DOLS in the table denote heterogeneous fully-modified OLS and heteroge-

[9] See Kao (1999) and Kao and Chiang (2000) for the consistent estimator of $\hat{\Omega}$. We use the Bartlett kernel with truncation lag length being 4 in estimating $\hat{\Omega}$. However, the following test results are unchanged if we use other lag lengths less or more than 4.

Table 3: Panel cointegration estimation

	OLS	FM	DOLS
Unemployment	0.560	0.552	0.623
	(20.313)	(16.402)	(14.241)
bias (%)		-1.482	10.092
Vacancies	0.302	0.258	0.289
	(12.398)	(7.655)	(6.606)
bias (%)	-	-17.195	-4.466
RTS	0.862	0.810	0.912
Wald test: constant RTS	10.791	95.811	12.066
(p-value)	(0.001)	(0.000)	(0.001)
R^2	0.968	0.942	0.802

Note: t-values are in parentheses. The lag length of the Bartlett kernel in estimating Ω_i is 4 for all i. For DOLS, we set lag = 3 and lead = 2. The reported bias is calculated by $(1 - \beta_{\mathrm{ols}}/\beta_j) \times 100$ (%), where β_j is replaced by each estimator.

neous Dynamic OLS, respectively. Their t-values are in parentheses. For comparison, conventional OLS estimates without bias-correction is also listed in the table[10].

Estimated coefficients of unemployment and vacancies, $\hat{\beta}_u$ and $\hat{\beta}_v$, exhibit correct signs and are statistically significant for all the cases. However, the values of estimates by different estimation methods differ from each other. The heterogeneous FM reduces both $\hat{\beta}_u$ and $\hat{\beta}_v$. Especially, the latter's reduction is considerable, approximately 17%. In the heterogeneous DOLS, $\hat{\beta}_u$ rises by approximately 10%,

[10] It is equivalent to the within estimates reported by Kano and Ohta (2002).

and $\hat{\beta}_v$ falls by approximately 4%[11].

The Monte Carlo experiment implemented by Kao and Chiang (2000) shows that the finite sample performance of DOLS greatly dominates those of alternatives[12]. Specifically, even in the case with shorter time dimension than in ours such as $T = 20$, the deviation of DOLS estimate from the true parameter value is surprisingly small. Therefore, following our estimation result of DOLS, we could claim that the OLS estimate of the hiring elasticity to unemployment is under-estimated, while that to vacancies is over-estimated.

The returns to scale of the matching function, RTS $= \beta_u + \beta_v$, is important in the equilibrium unemployment theory, since increasing RTS is consistent with multiple, Pareto-rankable "natural rates" of unemployment. In addition, constant RTS is often assumed for convenience. So testing constant RTS has been of particular interest for the proponents of the equilibrium unemployment theory. See the Wald test results in Table 3, which test H_0 : RTS $= 1$ against H_1 : RTS < 0[13]. It is shown that constant RTS is rejected for all cases, supporting decreasing RTS.

[11] We set the lag = 3 and lead = 4 for the DOLS. However, our estimates are not sensitive to the lag and lead length selected.

[12] Note that both FM and DOLS have the same asymptotic normal distribution. See Kao and Chiang (2000). They discuss that resulting poorer finite sample performance of the FM could be attributed to the failure of its non-parametric data transformation based on the estimated long-run covariance matrix.

[13] See Kao and Chiang (2000) for the Wald test for linear restrictions on the FM and the DOLS estimator.

4 Summary and conclusion

This paper investigated the long-run relationship among new hiring, unemployment and vacancies of Japanese labor market in the panel data framework.

Our main findings are as follows. First, new hiring, unemployment, and vacancies are $I(1)$ processes. Second, they are cointegrated, in other words, there exists a long-run relationship implied by the matching function in the theory of equilibrium unemployment. Third, conventional within estimates of hiring elasticity with respect to unemployment and vacancies have non-negligible biases. Finally, decreasing returns to scale of the matching function is statistically supported.

References

BLANCHARD, O. J., AND P. DIAMOND (1989): "The Beveridge Curve," *Brooking Papers on Economic Activity*, 1, 1–76.

GROSS, D. M. (1997): "Aggregate Job Matching and Returns to Scale in Germany," *Economics Letters*, 56, 243–248.

IM, K. S., M. H. PESARAN, AND Y. SHIN (1997): "Testing for Unit Roots in Heterogeneous Panels," working paper, University of Cambridge.

KANO, S., AND M. OHTA (2002): "Estimating a Matching Function and Regional Matching Efficiencies: Japanese Panel Data for 1973-1999," Institute of Policy and Planning Sciences, University of Tsukuba, Discussion Paper No. 997.

KAO, C. (1999): "Spurious Regression and Residual-Based Tests for Cointegration in Panel Data," *Journal of Econometrics*, 90, 1–44.

KAO, C., AND M.-H. CHIANG (2000): "On the Estimation and Inference of a Cointegration in Panel Data," working paper, Center for Policy Research, Syracuse University.

NG, S., AND P. PERRON (1995): "Unit Root Tests in ARMA Models with Data Dependent Methods for the Selection of the Truncation Lag," *Journal of the American Statistical Association*, 90, 268–281.

Papell, D. H., C. J. Murray, and H. Ghiblawi (2000): "The Structure of Unemployment," *Review of Economics and Statistics*, 82(2), 309–315.

Petrongolo, B., and C. A. Pissarides (2001): "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 36, 390–431.

Phillips, P. C. B., and H. Moon (1999): "Linear Regression Limit Theory for Nonstationary Panel Data," *Econometrica*, 67, 1057–1111.

Pissarides, C. A. (2000): *Equilibrium Unemployment Theory*. The MIT Press, Cambridge, MA, second edn.

Schwert, G. W. (1989): "Tests for Unit Roots: A Monte Carlo Investigation," *Journal of Business and Economic Statistics*, 7, 147–159.