Discussion Paper Series

RIEB

Kobe University

DP2021-03

Agglomeration Effects on Job Matching Efficiency: Evidence from Japan

Yudai HIGASHI

February 9, 2021

Research Institute for Economics and Business Administration
Kobe University
2-1 Rokkodai, Nada, Kobe 657-8501 JAPAN

Agglomeration effects on job matching efficiency: Evidence from Japan*

Yudai Higashi†

Graduate School of Humanities and Social Sciences, Okayama University

Research Institute for Economics and Business Administration, Kobe University

February 9, 2021

Abstract

This paper examines the effect of agglomeration on the regional efficiency of matching of job seekers and vacancies by using Japanese regional panel data. We find that a higher population density improves or at least does not deteriorate regional matching efficiency for most of the sample periods, suggesting that the benefit of agglomeration tends to be significant. However, the effect is significantly negative in 2011 when a serious earthquake and tsunami occurred in Japan, suggesting that the congestion effect is superior to the benefit of agglomeration when labor markets suffer from damages, such as those caused by natural disasters.

Keywords: Job search; Matching function; Agglomeration; Local labor market

JEL classification: J64; R11

* This research was supported by JSPS KAKENHI Grant Number JP20K22093.

† E-mail: yudai.higashi@okayama-u.ac.jp

1. Introduction

The heterogeneity of agglomeration levels across local labor markets can lead to regional disparities in the matching of job seekers and vacancies. Some studies, using microdata, show that more densely populated local labor markets increase job finding rates, suggesting that agglomeration makes the job matching process more efficient (de Blasio and Di Addario, 2005; Di Addario, 2011; Andini et al., 2013; Abel and Deitz, 2015).

Based on the evidence from microdata, we can consider that the job matching process at the regional aggregate level also seems to be more efficient in denser local labor markets. By estimating the aggregate matching functions, which explains the number of matches by the numbers of job seekers and vacancies, Coles and Smith (1996) and Hynninen and Lahtonen (2007) find a positive correlation between population density and regional matching efficiency, supporting the benefits of agglomeration. However, Kano and Ohta (2005) found a negative correlation, suggesting that the congestion effect is superior to the benefits of agglomeration. At the regional aggregate level, previous studies indicate the two conflicting effects of agglomeration.

This study aims to offer evidence that both positive and negative agglomeration effects on regional matching efficiency are supported. We assume that the balance of benefit and cost of agglomeration can vary based on the conditions of the labor market. To empirically analyze this conjecture, we examine whether the agglomeration effect changes over time by using Japanese regional panel data for 2008–2018. During this sample period, the conditions of the Japanese labor market were not constant because the country not only experienced a phase of economic recovery after the Great Recession, but also a catastrophic earthquake and tsunami in 2011. Such conditions may vary the agglomeration effect on regional matching efficiency.

The remainder of this paper is organized as follows. Section 2 presents the empirical methodology and data. Section 3 discusses the estimation results. Section 4 concludes the paper.

2. Methodology

2.1. Empirical model

To determine regional matching efficiency, we first estimate the matching function. Following previous empirical studies on matching functions (Petrongolo and Pissarides, 2001), the specification is assumed to be a Cobb–Douglass form, as follows:

$$M_{it} = A_{it} U_{it-1}^{\alpha} V_{it-1}^{\beta}, \tag{1}$$

where M_{it} denotes the number of outflows from unemployment in region i during period t. U_{it-1} and V_{it-1} are the stocks of the unemployed and vacancies in region i at the beginning of period t, respectively. α and β are their elasticities. The search and matching theory expects that their signs are positive (Pissarides, 2000). Finally, A_{it} represents the matching efficiency, given by $A_{it} = \exp(\psi_i + \varphi_t + \varepsilon_{it})$. That is, the matching efficiency is assumed to be decomposed into a time-invariant region-specific matching efficiency ψ_i, a time-specific matching efficiency φ_t, and a random shock ε_{it}. Using regional panel data, we estimate the log-linear specification of equation (1):

$$\ln M_{it} = \alpha \ln U_{it-1} + \beta \ln V_{it-1} + \psi_i + \varphi_t + \varepsilon_{it}. \tag{2}$$

In this equation, we can obtain regional matching efficiency as the region fixed effect ψ_i.

To estimate equation (2), we use Japanese regional panel data over 132 months from January 2008 to December 2018. Here, we focus on whether the agglomeration effect on matching efficiency changes over time. Therefore, we estimate equation (2) by yearly subsamples, covering 12 months from January to December in the corresponding year, to obtain regional matching efficiencies by year. That is, we assess yearly changes in regional matching efficiency.

Next, we present a model to estimate the agglomeration effect on regional matching efficiency. Following Hynninen and Lahtonen (2007) and Kano and Ohta (2005), the estimation model is given by

$$\hat{\psi}_i = \gamma + \delta \ln D_i + \xi_i, \tag{3}$$

where $\hat{\psi}_i$ is the regional matching efficiency estimated by equation (2). D_i is the population density in region i. δ, the interest parameter, captures the agglomeration effect on regional matching efficiency. γ is a constant, and ξ_i is an error term. The regional population data for D_i is available at the yearly level. Therefore, we regress the regional matching efficiency that is estimated by a subsample of 12 months in each year to population density in the corresponding year.

The ordinary least squares (OLS) estimations for equation (3) may yield results suffering from endogenous bias because job seekers may choose their locations based on the conditions of local labor markets. To address this problem, we also run two-stage least squares (2SLS) estimations by utilizing population density in 1920 as the IV.[1] This IV seems to be valid. This is because job seekers' behaviors in the sample periods (i.e., January 2008–December 2018) should be different from such past ones, indicating that the IV satisfies the exclusion restriction. Furthermore, population density in the sample periods corelate with past one because cities have gradually grown, indicating that the IV also satisfies the relevance.

[1] The population data in 1920 comes from the Population Census.

2.2. Data

We utilize two data sources. First, the data on the outflows from unemployment, the stock of unemployed, and the stock of vacancies for the matching functions come from the Report on Employment Service, provided by the Japanese Ministry of Health, Labour and Welfare. We use panel data for January 2008–December 2018 (i.e., 132 months) and 434 Public Employment Security Office (PESO) jurisdiction regions.

Second, to estimate the effect of agglomeration on regional matching efficiency, we use population data for 2008–2018 from the Basic Resident Registration, provided by the Ministry of Internal Affairs and Communications. This data contains the number of persons who are registered as residents at the municipal level by year.[2] To merge this data with the data for the matching functions, we aggregate the municipal-level population data at the PESO jurisdiction region level because each PESO covers some municipalities.[3] Then, we calculate population density as the number of residents per square kilometer of the inhabitable area.[4]

3. Results

3.1. Estimations of regional matching efficiency

To obtain regional matching efficiency, we first estimate the matching functions by yearly subsamples as well as whole sample periods.[5] Table 1 shows the correlations between the estimated regional matching efficiencies by year. The "overall" matching

---

[2] The population in this data is as of March 31 for 2008–2013 and as of January 1 for 2014–2018.

[3] Appendix A describes how to arrange regional units for the dataset in detail.

[4] Tables A1 and A2 shows the summary statistics of variables for the matching functions and population density, respectively.

[5] Table A3 shows the estimation results of the matching functions.

efficiencies are estimated by the matching function for whole sample periods, while each year's matching efficiency is estimated for the subsample of the corresponding year. We find that all correlation coefficients are positive.

[Table 1]

The magnitudes of correlations between 2011 and the other years are smaller. We consider that some structural changes occurred in the Japanese labor market in 2011, suggesting that the agglomeration effect on regional matching efficiency in this year could differ from that of other years. We examine this conjecture using regression analyses in the following subsection.

3.2. Agglomeration effect on regional matching efficiency

Table 2 shows the estimation results of the agglomeration effect on regional matching efficiency for the whole sample periods.[6] According to the OLS estimation results in column (1), the coefficient of log of population density is significantly positive. The 2SLS estimation confirms a similar result, as shown in column (2). These results suggest that agglomeration improves regional matching efficiency.

[Table 2]

Next, we consider whether the agglomeration effect on regional matching efficiency is significant over time. Although the null hypothesis that the variable is exogenous is not rejected for the 2SLS estimations for whole sample periods in column (2) of Table 2, hereafter, we show the 2SLS estimation results for each year's regional matching efficiency. This is because the test for some years shows endogeneity. Table 3 presents

[6] For the estimations for the whole sample periods, log of population density in 2008, the initial period over the sample periods, is adopted as the independent variable.

the 2SLS estimation results by year. We find different results across years. The effects of population density are significantly positive for 2008, 2012, 2014, and 2017, while they are not statistically significant for 2009, 2010, 2013, 2015, 2016, and 2018. Furthermore, the coefficient is significantly negative for 2011.

[Table 3]

On the whole, the agglomeration effect improves or at least does not deteriorate matching efficiency in most periods. That is, the benefit of agglomeration could generally be larger than or similar to the congestion effect. In contrast, the negative effect of population density in 2011 suggests that the congestion effect is larger than the benefit of agglomeration. In March of the year, the Great East Japan Earthquake occurred. This earthquake generated a tsunami and the accident at the Fukushima Daiichi Nuclear Power Station, resulting in serious damage to the Japanese labor market. After the earthquake, the compositions of the unemployed and vacancies changed (Higuchi et al., 2012). Then, matching efficiencies in the damaged regions and their surrounding regions were temporarily deteriorated (Higashi, 2020). We consider that such unusual changes in the labor market boost the congestion effect to the extent that it exceeds the benefit of agglomeration. Consequently, the job matching process in denser local labor markets could not be conducted efficiently. Overall, the agglomeration effect on regional job matching efficiency depends on the conditions of the labor market.

4. Conclusion

This study examined whether the agglomeration effects on regional job matching efficiency change over time. Using Japanese data, we found that the agglomeration effects differ across years. In most periods, the effects are significantly positive or insignificant.

In contrast, we found a significantly negative effect in 2011 when a serious earthquake and tsunami occurred. These results suggest that generally, the benefit of agglomeration tends to be superior or similar to the congestion effect in Japanese local labor markets. However, when labor markets suffer from damages, such as those caused by natural disasters, the congestion effect exceeds the benefit of agglomeration. We conclude that whether agglomeration improves regional matching efficiency depends on the conditions of the labor market.

Appendix A. Arrangement of regional units

Some PESO jurisdiction regions and municipalities were merged during the sample periods (i.e., January 2008–December 2018). Thus, we arrange their regional units as of 2015 when they had last merged. Some municipalities that were classified into wards in ordinance-designated cities separated between 2008 and 2012. The same was the case for periods between 1920 (i.e., the year for population as the IV) and 2008. In these cases, we distribute their population based on the areas of separated municipalities to arrange the regional unit as of 2015. For this calculation, we use the Municipality Map Maker for Web (http://www.tkirimura.com/mmm/), developed by Kirimura et al. (2011).

Appendix B. Tables

[Tables A1–A4]

References

Abel, J. R., and R. Deitz. 2015. "Agglomeration and Job Matching among College Graduates." *Regional Science and Urban Economics* 51: 14–24.

Andini, M., G. de Blasio, G. Duranton, and W. C. Strange. 2013. “Marshallian Labour Market Pooling: Evidence from Italy.” *Regional Science and Urban Economics* 43 (6): 1008–1022.

Coles, M. G., and E. Smith. 1996. “Cross-Section Estimation of the Matching Function: Evidence from England and Wales.” *Economica* 63 (252): 589.

de Blasio, G., and S. Di Addario. 2005. “Do Workers Benefit from Industrial Agglomeration?.” *Journal of Regional Science* 45 (4): 797–827.

Di Addario, S. 2011. “Job Search in Thick Markets.” *Journal of Urban Economics* 69 (3): 303–318.

Higashi, Y. 2020. “Effects of Region-Specific Shocks on Labor Market Tightness and Matching Efficiency: Evidence from the 2011 Tohoku Earthquake in Japan.” *Annals of Regional Science* 65 (1): 193–219.

Higuchi, Y., T. Inui, T. Hosoi, I. Takabe, and A. Kawakami. 2012. “The Impact of the Great East Japan Earthquake on the Labor Market: Need to Resolve the Employment Mismatch in the Disaster-Stricken Areas.” *Japan Labor Review* 9 (4): 4–21.

Hynninen, S.-M., and J. Lahtonen. 2007. “Does Population Density Matter in the Process of Matching Heterogeneous Job Seekers and Vacancies?” *Empirica* 34 (5): 397–410.

Kano, S., and M. Ohta. 2005. “Estimating a Matching Function and Regional Matching Efficiencies: Japanese Panel Data for 1973–1999.” *Japan and the World Economy* 17 (1): 25–41.

Kirimura, T., T. Nakaya, and K. Yano. 2011. “Building a Spatio-temporal GIS Database about Boundaries of Municipalities: Municipality Map Maker for Web.” *Theory and*

Applications of GIS 19 (2) 83–92 (in Japanese).

Petrongolo, B., and C. A. Pissarides. 2001. "Looking into the Black Box: A Survey of the Matching Function." *Journal of Economic Literature* 39 (2): 390–431.

Pissarides, C. A. (2000). *Equilibrium Unemployment Theory*. 2nd ed. Cambridge: MIT Press.

Table 1: Correlations between regional matching efficiencies across years

	Overall	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Overall	1											
2008	0.90	1										
2009	0.95	0.88	1									
2010	0.95	0.81	0.96	1								
2011	0.62	0.28	0.64	0.74	1							
2012	0.98	0.90	0.93	0.93	0.58	1						
2013	0.94	0.72	0.90	0.93	0.79	0.92	1					
2014	0.98	0.85	0.92	0.92	0.63	0.96	0.95	1				
2015	0.97	0.79	0.90	0.91	0.67	0.93	0.94	0.97	1			
2016	0.84	0.55	0.78	0.84	0.83	0.78	0.91	0.86	0.91	1		
2017	0.97	0.89	0.89	0.87	0.48	0.95	0.86	0.95	0.94	0.80	1	
2018	0.95	0.77	0.87	0.89	0.64	0.90	0.91	0.94	0.96	0.90	0.95	1

Table 2: Estimation results for agglomeration effect on regional matching efficiency for whole sample periods

	Dependent variable: Region matching efficiency		
	OLS	2SLS	
		Second stage	First stage
	(1)	(2)	(3)
ln(Population density)	0.045***	0.042**	
	(0.011)	(0.017)	
ln(Population density in 1920)			1.025***
			(0.053)
Observations	434	434	434
Adj. R2	0.046		0.494
First stage F statistic		367.281	
Robust score chi2		0.072	
p-value		0.788	

Notes: Robust standard errors are in parentheses. ***, **, and * represent statistical significance at 1%, 5%, and 10% level, respectively. Constant is not reported. The IV for 2SLS is the log of population density in 1920. Robust score chi2 statistic is for test of exogeneity.

Table 3: 2SLS estimation results for the agglomeration effect on regional matching efficiency by year

	Dependent variable: Region matching efficiency										
	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
ln(Population density)	0.132***	0.023	0.006	-0.054***	0.047***	0.008	0.039**	0.020	-0.017	0.058***	0.015
	(0.020)	(0.018)	(0.018)	(0.017)	(0.017)	(0.015)	(0.016)	(0.017)	(0.019)	(0.019)	(0.018)
Observations	434	434	434	434	434	434	434	434	434	434	434
First stage F statistic	367.281	368.852	370.184	371.479	372.562	383.782	384.958	386.415	387.923	389.326	390.863
Robust score chi2	4.342	1.131	1.200	2.678	1.432	0.621	0.011	0.029	0.606	0.856	0.289
p-value	0.037	0.288	0.273	0.102	0.231	0.431	0.917	0.864	0.436	0.355	0.591

Notes: Robust standard errors are in parentheses. ***, **, and * represent statistical significance at 1%, 5%, and 10% level, respectively. Constant is not reported. The IV for 2SLS is the log of population density in 1920. Robust score chi2 statistic is for test of exogeneity.

Table A1: Summary statistics of variables for the matching function

	Mean	Std. Dev.	Min	Max
Outflows from unemployment	362.47	247.05	20	1907
Stock of unemployed	4917.21	4399.88	242	32831
Stock of vacancies	4788.42	6287.34	249	63003

Notes: The number of observations is 57,288 across 434 PESO jurisdiction regions and 132 months from January 2008 to December 2018.

Table A2: Summary statistics of population density by year (/km2)

	Mean	Std. Dev.	Min	Max
2008	1933.56	2910.77	49.07	16292.24
2009	1938.08	2930.95	48.63	16413.95
2010	1940.63	2945.07	48.34	16565.97
2011	1941.43	2955.71	47.91	16707.12
2012	1940.59	2965.13	47.63	16868.35
2013	1982.98	3084.77	47.55	18388.13
2014	1986.06	3098.27	47.30	18574.88
2015	1987.69	3118.30	46.74	18773.00
2016	1990.71	3143.03	46.17	18966.27
2017	1993.65	3167.05	45.49	19171.32
2018	1996.16	3189.94	44.85	19400.00
1920	825.33	2464.69	21.87	43485.06

Notes: The number of observations is 434. Population densities as as of March 31 for 2008–2013 and as of January 1 for 2014–2018.

Table A3: Estimation results of the matching functions by year

	Dependent variable: lnM											
	Overall	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
lnU	0.402***	0.220***	0.288***	0.306***	0.462***	0.335***	0.445***	0.412***	0.413***	0.437***	0.381***	0.346***
	(0.008)	(0.033)	(0.036)	(0.038)	(0.058)	(0.037)	(0.050)	(0.038)	(0.041)	(0.051)	(0.043)	(0.051)
lnV	0.178***	0.148***	0.281***	0.311***	0.314***	0.215***	0.216***	0.192***	0.210***	0.264***	0.150***	0.252***
	(0.005)	(0.026)	(0.026)	(0.025)	(0.031)	(0.031)	(0.026)	(0.026)	(0.034)	(0.037)	(0.035)	(0.041)
Observations	57288	5208	5208	5208	5208	5208	5208	5208	5208	5208	5208	5208
Adj. R2	0.970	0.973	0.967	0.973	0.967	0.977	0.978	0.979	0.974	0.978	0.979	0.976

Notes: All models contain region and time fixed effects. Robust standard errors are in parentheses. ***, **, and * represent statistical significance at 1%, 5%, and 10% level, respectively. Constant is not reported.

Table A4: First stage of 2SLS estimation for the agglomeration effect on regional matching efficiency by year

	Instrumented variable: ln(Population density)										
	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
ln(Population density in 1920)	1.025***	1.028***	1.031***	1.034***	1.037***	1.048***	1.050***	1.053***	1.058***	1.062***	1.066***
	(0.053)	(0.054)	(0.054)	(0.054)	(0.054)	(0.053)	(0.053)	(0.054)	(0.054)	(0.054)	(0.054)
Observations	434	434	434	434	434	434	434	434	434	434	434
Adj. R2	0.494	0.493	0.491	0.490	0.488	0.489	0.488	0.487	0.486	0.484	0.483

Notes: Robust standard errors are in parentheses. ***, **, and * represent statistical significance at 1%, 5%, and 10% level, respectively. Constant is not reported.