

Lahtonen, Jukka Petteri; Hynninen, Sanna-Mari

**Conference Paper**

## Does population density matter in the matching process of heterogeneous job seekers and vacancies?

45th Congress of the European Regional Science Association: "Land Use and Water Management in a Sustainable Network Society", 23-27 August 2005, Amsterdam, The Netherlands

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

*Suggested Citation:* Lahtonen, Jukka Petteri; Hynninen, Sanna-Mari (2005) : Does population density matter in the matching process of heterogeneous job seekers and vacancies?, 45th Congress of the European Regional Science Association: "Land Use and Water Management in a Sustainable Network Society", 23-27 August 2005, Amsterdam, The Netherlands, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/117639>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

**45<sup>th</sup> Congress of the European Regional Science Association**

**23-27 August 2005, Vrije Universiteit Amsterdam**

**Does population density matter in the matching process of heterogeneous job seekers and vacancies?**

Sanna-Mari Hynninen<sup>a</sup>, Jukka Lahtonen<sup>b</sup>

*University of Jyväskylä, School of Business and Economics, Jyväskylä, Finland*

**Abstract**

This paper studies the matching process of job seekers and vacancies in a collection of local labour markets. We measure the differences in the ability of the local markets to form new matches and trace whether these differences can be explained by the differing population densities across markets or by the distributions of education level of job seekers. We find that on average high-density areas are more productive in forming matches than others are.

**Keywords:** Matching function, local labour markets, population density, education, heterogeneity

---

<sup>a</sup> E-mail: sanna-mari.hynninen@econ.jyu.fi.

<sup>b</sup> E-mail: jukka.lahtonen@econ.jyu.fi

# 1. Introduction

In the recent literature on labour market matching, increased emphasis has been placed on the heterogeneity of both job seekers and vacant jobs (Burgess 1993; Pissarides 1994; Anderson and Burgess 2000; Burgess and Turon 2003; van Ours 1995; Broersma and van Ours 1998; Mumford and Smith 1999; Fahr and Sunde 2001). Heterogeneity of job seekers and vacancies appears to be the main source of frictions in the labour market and therefore crucial subject when matching aspects of the labour market are investigated. In addition, many studies concern the labour market not as a whole but as a collection of several heterogeneous micro-markets.

This paper analyses the matching of vacant jobs and job seekers that differ in their education level in local labour markets that differ from each other by their population densities. The motivation of the study comes mainly from the three sources. (i) Coles and Smith (1996) argue that population density, not the numbers of vacant jobs and job seekers, determines the matching rates in the labour market. According to them, the matching process is the more effective the more concentrate the market is. That is because the communication of parties close to each other requires lower effort and costs<sup>1</sup>.

(ii) Kano and Ohta (2005) take another view on the role of the density. Apart from the idea of the lower effort and costs of communication of parties close to each other, there may be another factor affecting the matching process. Namely, the heterogeneity of job seekers and vacancies may be higher in densely populated areas than elsewhere<sup>2</sup>, which may increase search frictions. Successful matches may appear easily if the skill-distribution of workers and the hiring standards of firms are concentrated on the same level. But in urbanised areas, where the distributions tend to be wide, the concentration of characteristics of job seekers may differ from that required by employers. It im-

---

<sup>1</sup> They illustrate the argument by the following mind experiment: a blindfolded man and a woman wander around the field seeking to contact each other. Obviously, the contact rate of 1000 men and women wandering in the same size of field would be more than 1000 times the contact rate of the original situation of one man and woman.

<sup>2</sup> For example, Coles and Smith (1995) report several demographic variables significant at explaining matching rates. According to their results, the matching rates increase in towns that have younger population and decreases with population with high education. Their matching process exhibits constant returns to scale. Thus, the results indicate that it is density that matters, not the numbers of job seekers and open vacancies in the labour market. However, they conclude that successful matches may have better quality in large cities than elsewhere since wages tend to be higher in large cities than elsewhere. There are larger markets for highly specialised workers in large cities.

plies difficulties in the matching process despite that the actors were located near each other in a geographical sense.

(iii) Wahba and Zenou (2003) consider population density as a proxy for the size of social networks and therefore also for the speed of information transmission. They conclude that as long as network remains a reasonable size, the size has a positive effect on matching efficiency. However the effect may become negative in very densely populated areas because of the dominating opposite congestion effect.

In this paper, we estimate the abilities of Local Labour Offices (LLOs) in Finland to form successful matches at given levels of inputs in the matching process. In addition, we investigate if the total factor productivity of the matching process deviates in areas with deviant population density. By controlling for the heterogeneity in the labour market allowing different employability for job seekers with different education levels we can estimate matching productivity that do not result from distributions of the education level of job seekers and the education requirements of vacancies.

Our data are informative and temporally, spatially, as well as by education level of job seekers and education requirements of open vacancies highly disaggregated. It consists of monthly panel from 146 LLOs in Finland over 10 years. Job seekers are divided in three groups by their education levels. These groups are primary, secondary, and high education. The data set is collected from the registers of the Ministry of Labour. Therefore, it includes job seekers and vacancies registered at the public employment agency which has an important role in the labour market in Finland.<sup>3</sup>

The paper is organised as follows: Section 2 defines a concept productivity of the matching process. A brief discussion about the estimation technique is also conducted. Section 3 introduces the model, section 4 describes the data, and in section 5 the results are presented. Finally, section 6 concludes the paper.

---

<sup>3</sup> The proportion of jobs mediated by LLOs in Finland is quite high. It varied between 1993 and 2002 from 49 per cent to 71 per cent being lowest in 1993 and highest in 1996 (Hämäläinen 2003). On average, the share is about 60 per cent. Public employers have a statutory duty to report an open vacancy, whereas for private firms this is optional. Despite this, the largest reported share of open vacancies is in the private sector (Räisänen 2004).

## 2. Productivity of the matching process

The total factor productivity of the matching process captures the ability of the labour market to form successful matches by given job seekers and vacancies. It is partly determined by a product of the rate at which job seekers and employers meet and the probability that a contact leads to a successful match (Anderson and Burgess 2000). The total factor productivity is defined as a fraction of production output to production factors in a standard production theory. We estimate the total factor productivity of the matching process and focus on differences in it between areas with deviant population density. By controlling for the educational structure of the stock of job seekers we clarify if the heterogeneity of job seekers affects productivity of the matching process.

In the case of a Cobb-Douglas matching function:

$$M_{i,t} = A_{i,t} U_{i,t-1}^{\alpha} V_{i,t-1}^{\beta} = A e^{\mu_i + \lambda_t + \varepsilon_{i,t}} U_{i,t-1}^{\alpha} V_{i,t-1}^{\beta}, \quad (1)$$

it is natural to interpret the parameter  $A_{i,t}$  to denote the productivity part of the function since it is independent of  $U$  and  $V$  but may vary across LLOs (part  $\mu_i$ ), and over time (part  $\lambda_t$ ). The error term is assumed to be identically and independently distributed with expectation 0 and constant variance. If we furthermore impose additional assumption that

$$\sum_{i=1}^n \mu_i = \sum_{t=1}^T \lambda_t = 0, \quad (2)$$

the equation becomes the fixed-effects model which can be estimated by the Least Squares Dummy Variables estimator, or equivalently by the within estimator. This estimator is, however, known to suffer some problems; see for example Ibourk et al. (2004). There is potential correlation between fixed effects and other explanatory variables, which might imply an estimation bias and catch the return-to-scale effects.

Our data is comprised of LLOs in which the proportion of vacant jobs to job seekers is on average very low (2 - 3 percent). In other words, the number of job seekers is high meaning that after the log-transformation it shows little time-variation within the LLOs; see Figure 1. There is, however, some variation between LLOs. These features of the data are confirmed by the estimation results; see Table 3. When the fixed-effect estimator is applied (Specifications 2-4) the coefficient of job seekers is remarkably lower than that estimated by random-effect estimator (Specification 1). The value given by random-effect estimator is higher because it exploits both between and within - variations, but is inconsistent when random effects are correlated with explanatory variables, which is the case here.

Bleakley and Fuhrer (1997) and Wall and Zoega (2002) estimate shifts in the efficiency of the matching process interpreting region-specific fixed-effects as efficiency terms. Being aware of problems with interpretations of the term efficiency in the context of the fixed-effects model (see i.e. Ibourk et al. 2004) we estimate the average productivity of the matching process in areas with differing population density. We control for the educational structure of job seekers to find out if it affects matches in LLOs and to find out if it explains part of the productivity. After controlling for the educational structure, estimated fixed effects indicate total factor productivity that is not owing to the educational structure.

It is worth noting that education of job seekers, per se, does not make matching process more productive. It calls for successful matching between education level of job seekers and education requirements of vacancies that the matching works more productively. Therefore, employability differences cannot directly be derived from search intensity or ranking behaviour of firms but rather from matching between supply and demand of education. Thus, referring to Coles and Smith (1996), Kano and Ohta (2005), and Wahba and Zenou (2003) population density of the area connected with education structure of job seekers is interesting to investigate. We assume that owing to the interactive nature of the production of matches, controlling for the heterogeneity of job seekers also controls the requirements of the vacancy side. There are no positive effects due to a particular educational level of a job seeker on the employability if there is no demand for that kind of education.

### 3. Model

In a basic matching function the inputs are the number of job seekers  $U$  and the number of vacant jobs  $V$ . The output is the job matches, i.e. filled vacancies  $M$  formed within a given time period. If job seekers or firms search more intensely for a match, we observe the number of matches to increase, like the matching technology or efficiency of it had been improved. Thus, denote the average search intensity, or employability for now on, of job seekers by  $s$  and the corresponding employability of vacancies by  $a$ . The aggregate matching function is now

$$M = m(sU, aV) \quad (3)$$

where the first argument,  $sU$ , is called the efficiency units of job seekers and the latter argument is the efficiency units of vacancies. Furthermore, the pools of job seekers and vacancies can be distinguished into three groups according to their educational level and educational requirement level, respectively. Note that now  $s$  is the average of group-specific employability of job seekers and  $a$  the average ability of vacancies to become filled.

$$M = m\left(\sum_{i=1}^3 s_i U_i, \sum_{j=1}^3 a_j V_j\right) = m(sU, aV) \quad (4)$$

For estimation purposes, we can choose to measure employability of different groups with respect to employability of one particular group, let it be  $s_2$  in case of job seekers. In the model, employability of these groups is set to equal to 1. Due to lack of the data, we are not able to allow different employability for vacancies with different education requirements. Our modelling follows the appendix of Ibourk et al. (2004) taking however somewhat different view on modelling heterogeneity in the matching model. By adding and subtracting terms the function comes into the following Cobb-Douglas form (See Appendix for more details):

$$M_{i,t} = A \left( (U + (s_1 - 1)U_1 + (s_3 - 1)U_3)_{i,t-1}^\alpha V_{i,t-1}^\beta \right) \quad (5)$$

Dividing the job seeker input by  $U$  and taking logs yields

$$\ln M_{i,t} = c_i + \alpha \ln(U)_{i,t-1} + \alpha \ln(1 + \gamma(\frac{U_1}{U})_{i,t-1} + \eta(\frac{U_3}{U})_{i,t-1}) + \beta \ln(V)_{i,t-1}, \quad (6)$$

where  $\gamma$  denotes  $(s_l - l)$  and  $\eta = (s_3 - l)$  respectively. Taylor-expansion provides linear approximation for the last term of the expression.

$$\ln M_{i,t} = c_i + \alpha \ln(U)_{i,t-1} + \alpha \gamma(\frac{U_1}{U})_{i,t-1} + \alpha \eta(\frac{U_3}{U})_{i,t-1} + \beta \ln(V)_{i,t-1} \quad (7)$$

The parameters of the expression have clear interpretations,  $\alpha$  tells the overall elasticity of matches with respect to job seekers and  $\beta$  that for vacancies. The coefficients for share variables tell the average percentage change of matches given the change of one percentage point in the relative share of that education group. Further, we can calculate from the estimates the employability of other groups with respect to the group of secondary educated job seekers. Therefore, both the overall elasticity of matches with respect to inputs and differences in employability formulate the significance of the composition of stocks of job seekers and vacancies in the matching process. On the other hand, the composition of stocks partly determines the elasticity of matches with respect to stocks.

#### 4. Data description

We have the time series of 146 local labour offices, LLOs, in Finland. The source of the data is the unemployment register of the Ministry of Labour. Each series contain observations from the period 1995:1 - 2004:9. The variables are filled vacancies within a month, the number of vacant jobs at the end of each month and the number of job seekers at the end of each month. The last-mentioned variable is furthermore distinguished between three levels of education, namely primary, secondary



and high level of education<sup>4</sup>. Table 1 provides descriptive statistics for the variables at the LLO-level.

The distributions of the variables show that for all the variables the mean is clearly higher than the median. In addition, the maximums are in a different league from the other statistics. The statistics represents the fact that the majority of LLOs in Finland are rather small in size, but some LLOs are very large compared to the average. This suggests the modelling using logarithmic transformation of the data. There are also LLOs in which the number of filled vacancies or vacant jobs is zero within the observation period. To avoid missing observations we add one to all observations before the log-transformation is applied.

In order to study the role of population density on matching productivity we order LLOs on the basis of their average population densities (population/km<sup>2</sup>) during 1991-2001. Despite some minor changes in the population density within offices, the reciprocal order of LLOs has remained stable during the research period. Three groups of LLOs are formed by setting group one to include 36 LLOs with highest population density, and group three 36 LLOs with lowest population density. Thus, the groups consist of LLOs belonging to the first and fourth quartiles of the population density distribution. The rest of the offices belong to group two. Thus we have decided to consider the offices being situated at the tails of the population density distribution as a separate LLO-groups.

Figure 1 plots the time series of the average values of key variables. In the upper left-hand side picture the average is taken over all LLOs. The right-hand-side picture plots averages over LLOs belonging to group 1 and so on. The log-transformation is applied before plotting. The strong time variation is observed for both vacancy series, but not for the series of job seekers. This is, however, mainly the product of log-transformation. The number of filled vacancies, contrary to the theory assumptions, seems to exceed the number of vacant jobs. The obvious explanation is that a major

---

<sup>4</sup> In the original data set there were nine distinguished levels of education. This system of classification has been used by Statistics Finland since 1971, and has been revised in 1997 to correspond to the international standard ISDEC 1997. For the estimations we aggregate the job seekers into groups of three education levels. Table 2 presents the relation between the three-group-classification and ISDEC 1997.

part of vacancies announced in the labour office will be filled within a week or two, thus they become filled before they are registered in a pool of vacant jobs. Our data on different groups of job seekers does not, however, allow us to deal with flows in the matching model.

Applying to differences between the groups with differing population density, the time series of group 1 lies above the series of other groups showing the difference the amount of both inputs and output in the matching process. Variation in logged job seekers and its contribution to variation in logged filled vacancies is small in all groups, while variations in vacancies contribute to variation in successful matches clearly. It seems that the connection is strongest in the group of highest population density and weakest among group of lowest population density.

## **5. Estimation results**

### **5.1 Fixed- and random-effects models**

Fahr and Sunde (2001) estimate distinct matching functions for distinct markets, for job seekers with different education levels among others. They find that an additional job seeker in a group of lowest educational level creates a new match with relatively higher probability than in other groups. We estimate the effects of the educational structure both in the whole data and in the groups formed according to the population density. Then we put the density groups in the same model utilising linear mixed model technique and investigate the importance and the significance of the density grouping in the matching model (Verbege and Molenberghs 2001).

In order to get the overall picture of the matching function we also report the matching estimations for the whole data with 146 LLOs. The estimation method is the Prais-Winsten regression with panel corrected standard errors (see StataCorp. 2001). Disturbances are assumed to be heteroscedastic, i.e. each panel is allowed to have its own variance. In order to handle possible dependencies in matching processes across panels, disturbances are assumed to be contemporaneously correlated, i.e. each pair of panels is allowed to have its own covariance. It is also assumed that there is first-

order autocorrelation within panels and that the coefficient of AR(1) process is common to all the panels. The results without assumption of autocorrelation, heteroscedasticity, and contemporaneous correlation are also reported on the basic specification for the whole data. Durbin-Watson coefficient 1.6 indicates, however, that autocorrelation is useful to take into account, and the test for groupwise heteroscedasticity shows that it is extremely important to allow the panels to have their own variances (Table 3., Specification 2).

The coefficient for job seekers, 0.34, in the matching model is lower than that for vacancies, 0.42, which is common in the models with filled vacancies as a dependent variable (Table 3, specification 4). Positive externalities due to increases in inputs are not in either side large enough to cancel out negative congestion effects owing to a rise in inputs, and the matching process exhibits decreasing returns. With regard to the education level of job seekers, both highly and primary educated job seekers improve the production of successful matches in LLOs in relation to secondary educated. A percentage point increase in the share of highly educated job seekers increases matches by 2.3 per cent and a corresponding rise in the share of primary educated by 1.8 per cent. These coefficients indicate that there is relatively higher demand for primary and highly educated workers than for secondary educated.

To clarify the total factor productivity of the matching process, i.e. the ability of a labour market to produce matches at given stocks of job seekers and vacancies, we calculate the LLO specific fixed effects from the specification 4 with all LLOs (Table 3). The average value of all fixed effects is – 1.5, which means that independently of variations in the stocks of job seekers and vacancies, the number of produced successful matches is 0.22, i.e. the technology parameter  $A$  in the matching model is 0.22. The calculation of the average values of the fixed effects for LLOs that belong to the group of high population density shows that the average productivity for that group is 0.25, for the middle group 0.22, and for the group of low population density 0.21. Therefore, according this specification, the group of highest population density has the highest total factor productivity in the matching process and the group of lowest population density lowest total factor productivity.

## **Distinct estimations for density groups**

Next we estimate distinct matching models for distinct population density groups. According to the estimation results, in the group of highest population density, the matching process seems to work independently of changes in the stock of job seekers (Table 4). The stock of job seekers is so large that the congestion effects among job seekers cancel out all positive externalities due to increases in the stocks. The coefficient for vacancies is quite high, 0.46. Changes in vacancies and dummies for LLOs, months, and years explain changes in matches very well,  $R^2$  is as high as 0.9. It is possible that the share of vacancies filled by job seekers not registered in LLOs is larger in densely populated areas than elsewhere. New vacancies produce successful matches with job seekers not registered in LLOs, and the congestion among job seekers in LLOs remains strong. The market share of public employment agency in the job market tends also be lower in more urbanised areas than in periphery (Hämäläinen 2004), which also complicates the job searching process of registered job seekers.

In the group of lowest population density, the coefficient for job seekers is high, 0.52. It is higher than the coefficient for vacancies, which is 0.34. A rise in the share of primary or highly educated also improves matches in these areas. The pools of job seekers and vacancies as well as the number of filled vacancies in these areas are very small compared to the group of highest population density. In addition, unemployment rates are usually high. As a result of these details, vacancies reported in LLOs are more often filled by job seekers searching jobs through LLOs. In addition, due to inadequately functioning labour markets, the pool of job seekers is more stagnant than elsewhere, and also small changes in it positively affect the production of new matches. Active labour market programmes might also cause dynamics in the pool of job seekers more than in the group of high population density.

In the middle group, the importance of additional job seekers in the production of matches is zero like in the group of high population density. In the specification without education variables, the coefficient for job seekers is however nearly significant. The explanation might be that the coefficient for primary educated captures the effect of all job seekers since the number of them increases

in accordance with the number of job seekers. They constitute the largest group of job seekers. Highly educated job seekers seem not to have any significant effect.

Since it is somewhat misleading to compare results between separate estimations, we estimate next the model where all density groups are included in the same model. Then we are also able to investigate the statistical significance of the population density grouping.

## 5.2 Mixed-effects model

Next we employ a linear mixed model that serves additional information in connection with two aspects. First, we are able to estimate the coefficients for all density groups at the same time, which again allows statistical testing of possible differences. Secondly, we can relax the assumption of the same coefficients for vacant jobs and job seekers across LLOs. We are, after all, interested in the differences across LLOs.

The model may be expressed in a hierarchical, two-stage analysis form, or equivalently in the Laird-Ware form (Laird and Ware, 1982). Following the latter view the model can be expressed as:

$$\begin{aligned} y_{it} &= \beta_1 x_{1it} + \dots + \beta_p x_{pit} + b_{i1} x_{1it} + \dots + b_{iq} x_{qit} + \varepsilon_{it} \\ b_{ik} &\sim N(0, \psi_k^2), \text{Cov}(b_k, b_{k'}) = \psi_{kk'} \\ \varepsilon_{it} &\sim N(0, \sigma^2 \lambda_{itt}), \text{Cov}(\varepsilon_{it}, \varepsilon_{it'}) = \sigma^2 \lambda_{itt'} \end{aligned} \quad (8)$$

where  $y_{it}$  is the value of the response variable (filled vacancies) for the  $t$ th observation in the  $i$ th of 146 LLOs;  $\beta_1 \dots \beta_p$  are the fixed-effect coefficients identical for all LLOs (intercept and coefficients for vacant jobs, all job seekers and relative shares of education groups);  $b_{i1}, \dots, b_{iq}$  are the random-effects coefficients for LLOs, assumed to be multivariately normally distributed with covariance-variance structure  $\psi_{kk'}$  that is constant across LLOs.  $\varepsilon_{it}$  is normally distributed error term with specified covariance between LLOs. We expect random effects to be possibly correlated with explanatory variables, which is a common feature in econometric models. Therefore we do not inference

strong causal statements from the results, instead, we estimate effects that explanatory variables have on response variables according to the data.

We fit a model to the data including fixed effects for intercepts, vacant jobs, all job seekers and relative shares of education groups. We include random intercept and slopes. The differences between density groups are estimated using dummy-variables. Thus, the estimated model is:

$$\begin{aligned} \ln M_{i,t} = & c + cD_1 + cD_2 + \alpha \ln(U_{i,t-1}) + \alpha D_1 \ln(U_{i,t-1}) + \alpha D_2 \ln(U_{i,t-1}) + \\ & \alpha \gamma \left( \frac{U_1}{U} \right)_{i,t-1} + \alpha \gamma D_1 \left( \frac{U_1}{U} \right)_{i,t-1} + \alpha \gamma D_2 \left( \frac{U_1}{U} \right)_{i,t-1} + \alpha \eta \left( \frac{U_3}{U} \right)_{i,t-1} + \alpha \eta D_1 \left( \frac{U_3}{U} \right)_{i,t-1} + \alpha \eta D_2 \left( \frac{U_3}{U} \right)_{i,t-1} \quad (9) \\ & + \beta \ln(V_{i,t-1}) + \beta D_1 \ln(V_{i,t-1}) + \beta D_2 \ln(V_{i,t-1}) + \\ & b_{i1} + b_{i2} \ln(U_{i,t-1}) + b_{i3} \ln(V_{i,t-1}) + b_{i4} \left( \frac{U_1}{U} \right)_{i,t-1} + b_{i5} \left( \frac{U_3}{U} \right)_{i,t-1} + \varepsilon_{it} \end{aligned}$$

In above formulation,  $\alpha$ ,  $\beta$  and  $c$  are fixed effects common for all LLOs.  $D_1$  and  $D_2$  are dummy-variables representing LLOs belonging to the density group 1 and group 2, respectively.  $U_{i,t-1}$  denotes the number of job seekers in LLO  $i$  at period  $t-1$ , correspondingly  $V_{i,t-1}$  is the number of vacant jobs. The second line of the equation includes the proportion of primary/highly educated job seekers of total number of seekers. The index 1 indicates primary education and 3 high level of education. Finally,  $b_{i1}, \dots, b_{i5}$  are the random effect coefficients for LLO  $i$ , they are considered as random variables, not as parameters, so only standard errors of them are estimated.

The results from the restricted maximum likelihood estimation are reported in Table 5. Specification (5) does not include education variables. Both intercepts and slopes are allowed to vary across LLOs. Among low-density LLOs the average value of intercepts is -2.6, the coefficient for job seekers 0.7 and that for vacancies 0.4. Thus, there appears to be increasing returns to scale in that group. The interactions between explanatory variables and group dummies indicate that the intercept is 1.85 higher in high-density group and 2.3 higher in middle-group than in low-density group. The result indicates that factors of total productivity not related to changes in the numbers of job seekers and vacancies are the highest in the middle group and the lowest in the low-density group. The comparison of coefficients for job seekers reveals that congestion effects are stronger in the

middle and high-density groups than among low density LLOs. The coefficients for vacancies show opposite behaviour.

In specification (6) the effect of education distribution across job seekers is controlled by including education share variables to the model. Among low-density LLOs an increase in the relative share of primary educated job seekers has positive effect on matches whereas the corresponding share of highly educated seekers have negative effect. The corresponding coefficients for middle group do not deviate statistically from those estimated for the low-density group. Instead, among high-density LLOs both high and low education groups have lower estimated coefficients than in the low-density group.

This gives the impression that in group 3 the requirements of firms are concentrated on lower educated workers, and the increase in the relative number of highly educated would decrease the matching productivity. However, the value of the intercept decreases after the insertion of education variables, so the education variables rather explain the efficiency of those LLOs than the inefficiency. That is not the case among high-density LLOs. The insertion of education variables increases the intercept term, so they catch a part of the inefficiency among those LLOs. The results are very much in line with the starting hypothesis, that high-density LLOs are more productive in matching job seekers and jobs but wider distributions of job seekers' education levels and employers' education requirements in those LLOs cause some inefficiency.

## **6. Conclusions**

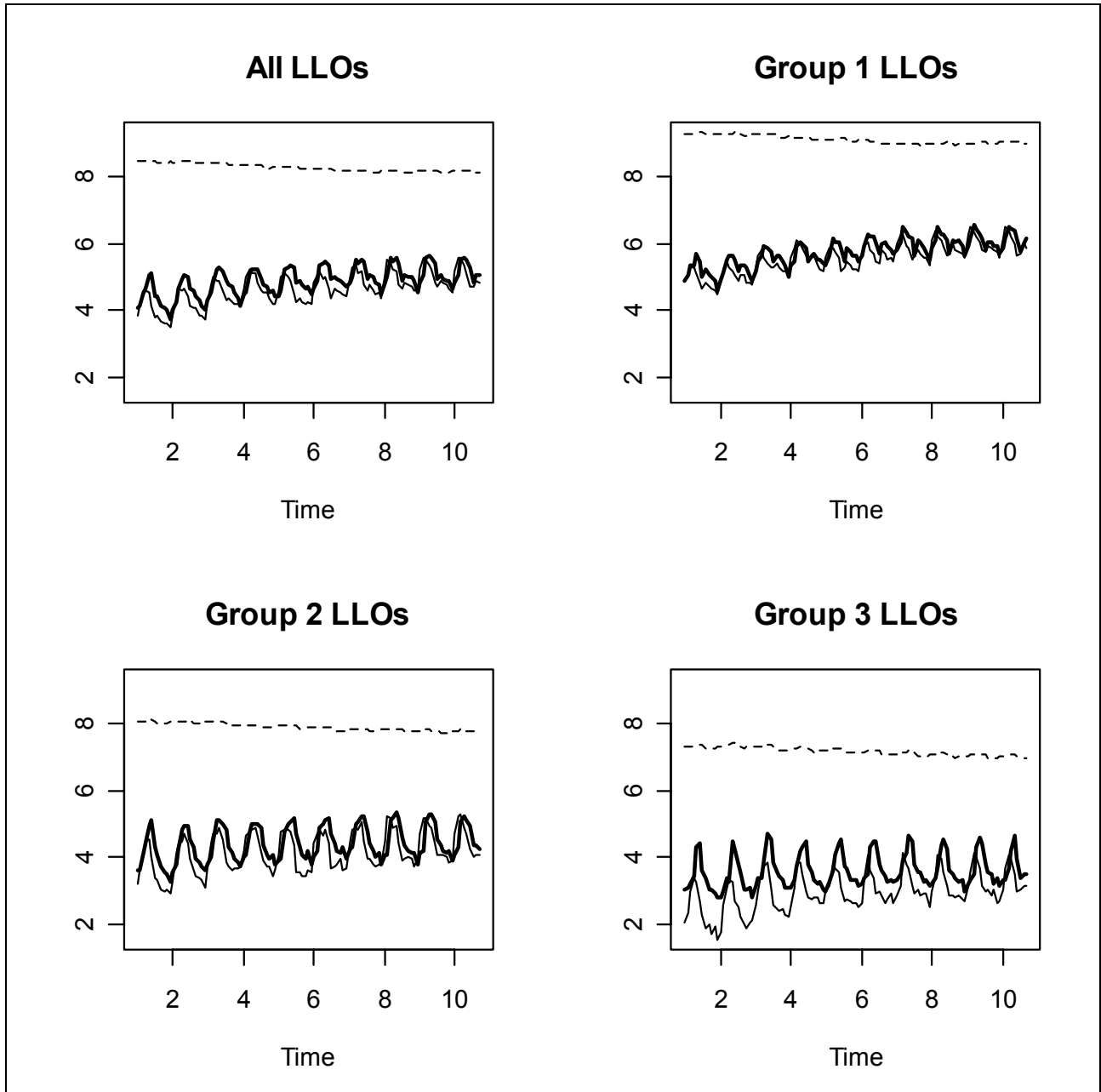
This paper studied the matching process of job seekers and vacant jobs in local labour markets in Finland. We estimated the ability of the local markets to form new matches and traced whether the differences in it could be explained by the differing population density across markets and by the distribution of the education level of job seekers. The starting hypothesis was that high-density areas are otherwise more productive than others but the productivity of them may be negatively affected by the heterogeneity of job seekers. Our results are in line with the hypothesis: after controlling the effect of average education level of job seekers, high-density areas are more productive in producing successful matches than others. The heterogeneity of job seekers seems to contribute to

the inefficiency of the matching process in high-density-areas more than elsewhere. Particularly, there seems to be more lower educated job seekers in high-density areas than would be necessary considering the requirements of employers. Therefore, we conclude that the wider heterogeneity of job seekers seems to cause frictions in the matching process in densely populated areas.



## Figures and tables

**Figure 1. Job seekers, filled vacancies and vacant jobs. Time series of averages taken over LLOs.**



Notes: The line ----- represents job seekers; — filled vacancies; — vacant jobs. All the variables are measured in logs. X-axis refers to time spanning from 1995:1 to 2004:9.

**Table 1. Descriptive statistics**

	<b>Filled Vacancies</b>	<b>Vacant Jobs</b>	<b>All Job Seekers</b>	<b>Primary Educated</b>	<b>Secondary Educated</b>	<b>Highly Educated</b>
<b>Min.</b>	0	0	183	83	65	6
<b>1st Qu.</b>	25	14	1225	630	484	75
<b>Median</b>	54	36	2082	1025	887	158
<b>Mean</b>	142	114	4045	1862	1725	457
<b>3rd Qu.</b>	120	91	3723	1830	1604	335
<b>Max.</b>	7717	7566	106329	49937	41946	20731

Notes: Statistics are calculated from monthly observations across all LLOs.

**Table 2. Relation between 3-group classification and ISDEC 1997.**

<b>ISDEC 1997</b>	<b>Name</b>	<b>3-group classification</b>
Level 0	Pre-primary education	-
Level 1	Primary education	1 Primary
Level 2	Lower secondary education	1 Primary
Level 3	Upper secondary education	2 Secondary
Level 4	Post secondary non-tertiary ed.	2 Secondary
Level 5	1st. stage of tertiary education: 5B-programmes	3 Highly
	5A-programmes	3 Highly
Level 6	2nd stage of tertiary education	3 Highly

**Table 3. Estimation results for random- and fixed-effects models**

<b>Variables</b>	<b>Specification</b>			
Dependent variable: $\ln M_t$	(1)	(2)	(3)	(4)
$\ln U_{t-1}$	0.41*** (0.023)	0.25*** (0.048)	0.32*** (0.081)	0.34** (0.077)
$\ln V_{t-1}$	0.43*** (0.005)	0.43*** (0.005)	0.42*** (0.008)	0.42*** (0.008)
<b>Education variables</b>				
$(HIGH/U)_{t-1}$				2.3*** (0.588)
$(LOW/U)_{t-1}$				1.8*** (0.295)
Constant	-1.05*** (0.183)			
Autocorrelation coefficient			0.2	0.19
Returns to scale	0.84***	0.68***	0.74***	0.76***
$R^2$	0.78	0.78	0.79	0.79
Number of observations	17 082	17 082	17 082	17 082
LR-test for groupwise heteroskedasticity		28 680.19***		
Durbin-Watson statistic		1.6		
<b>The average value of fixed effects</b>				
All				-1.5
Group 1				-1.39
Group 2				-1.52
Group 3				-1.57

Note: All models include yearly and monthly dummies. Models 2-4 include LLO-specific fixed effects. Model 1 is a random-effects model. Standard errors (in specifications 3-4 panel-corrected) are in parentheses. In specifications 2-4 the error terms are assumed to be contemporaneously correlated and autocorrelated AR(1) with autocorrelation coefficient common to all panel. \*\*\* denote statistical significance at the 0.1% level, \*\* at the 1% level, and \* at the 5% level. In tests for returns to scale, \*\*\* denote deviation from unity at the 0.1% level.

**Table 4. Estimation results for fixed-effects models conducted separately for three groups of LLOs**

<b>Variables</b>	<b>Group 1</b>		<b>Group 2</b>		<b>Group 3</b>	
	<b>Specification</b>		<b>Specification</b>		<b>Specification</b>	
Dependent variable: $\ln M_t$	(3)	(4)	(3)	(4)	(3)	(4)
$\ln U_{t-1}$	0.03 (0.077)	0.03 (0.077)	0.15 (0.087)	0.12 (0.086)	0.53*** (0.145)	0.51*** (0.145)
$\ln V_{t-1}$	0.46*** (0.013)	0.46*** (0.013)	0.44*** (0.01)	0.44*** (0.01)	0.34*** (0.012)	0.34*** (0.012)
<b>Education variables</b>						
$(HIGH/U)_{t-1}$		0.03 (0.651)		-0.1 (0.584)		2.15*** (1.022)
$(LOW/U)_{t-1}$		0.41 (0.409)		1.5*** (0.336)		1.43*** (0.478)
Autocorrelation coefficient	0.15	0.15	0.18	0.18	0.14	0.13
Returns to scale	0.46***	0.46***	0.44***	0.44***	0.87***	0.85***
$R^2$	0.9	0.9	0.73	0.73	0.63	0.63
Number of observations	4 212	4 212	8 658	8 658	4 212	4 212

Note: Group 1 includes the high population density LLOs, group 3 denotes the group of the low density, and group 2 is the middle group. All specifications include yearly and monthly dummies. Panel-corrected standard errors are in parentheses. The error terms are assumed to be contemporaneously correlated and autocorrelated AR(1) with autocorrelation coefficient common to all panel. \*\*\* denote statistical significance at the 0.1% level, \*\* at the 1% level, and \* at the 5% level. In tests for returns to scale, \*\*\* denote deviation from unity at the 0.1% level.

**Table 5. Estimation results for mixed-effects model**

Variables	Specification	
Dep. variable: $\ln M_t$	(5)	(6)
Intercept	-2.63*** (0.45)	-4.39*** (1.07)
$\ln U_{t-1}$	0.67*** (0.06)	0.76*** (0.11)
$\ln V_{t-1}$	0.40*** (0.02)	0.39*** (0.02)
Education variables		
$(HIGH/U)_{t-1}$		-0.87* (0.43)
$(LOW/U)_{t-1}$		1.82** (0.56)
<b>Interactions with d1:</b>		
d1*Intercept	1.85** (0.62)	4.73** (1.54)
d1* $\ln U_{t-1}$	-0.34*** (0.08)	-0.56*** (0.16)
d1* $\ln V_{t-1}$	0.09** (0.03)	0.10*** (0.03)
Education variables		
d1*(HIGH/U) $_{t-1}$		-1.01 (0.54)
d1*(LOW/U) $_{t-1}$		-2.00* (0.83)
<b>Interactions with d2:</b>		
d2*Intercept	2.30*** (0.56)	3.83** (1.32)
d2* $\ln U_{t-1}$	-0.37*** (0.07)	-0.51*** (0.14)
d2* $\ln V_{t-1}$	0.07** (0.02)	0.07** (0.02)
Education variables		
d2*(HIGH/U) $_{t-1}$		-0.60 (0.49)
d2*(LOW/U) $_{t-1}$		-0.82 (0.69)
<b>Random effects:</b>		
Intercept	0.32	5.29
$\ln U_{t-1}$	0.1	0.51
$\ln V_{t-1}$	0.04	0.09
$(HIGH/U)_{t-1}$		0.27
$(LOW/U)_{t-1}$		2.53
residual	0.44	0.44

Notes: Although not reported here, time trend and monthly dummies are included in both specifications to capture time effects. Significance levels: \*\*\* significant at level 0.1 %, \*\* at level 1%, \* at level 5 %. Dummy variable d1 refers to group of high density LLOs, d2 to middle group and low- density LLOs are the reference group.

## APPENDIX

An efficient stock of job seekers takes the following form:

$$ESS = (U_2 + s_1 U_1 + s_3 U_3) \quad (A.1)$$

The employability of group 2, secondary educated job seekers, is set equal to 1. The coefficient  $s_1$  denotes the employability of primary educated job seekers with respect to secondary educated, and  $s_3$  that of highly educated respectively. It also holds that

$$ESS = (U_1 + U_2 + U_3 + s_1 U_1 + s_3 U_3 - U_1 - U_3) \quad (A.2)$$

which is equal to

$$ESS = (U_2 + (s_1 - 1)U_1 + (s_3 - 1)U_3) \quad (A.3)$$

where  $U$  denotes the whole stock of job seekers. Thus, the matching function with the efficient stock of job seekers takes the form

$$M = A(U + \gamma U_1 + \eta U_3)^\alpha V^\beta \quad (A.4)$$

where  $\gamma$  denotes  $(s_1 - 1)$  and  $\eta$   $(s_3 - 1)$ . Dividing the job seeker input by  $U$  yields

$$M = A[U(1 + \gamma(\frac{U_1}{U}) + \eta(\frac{U_3}{U}))]^\alpha V^\beta \quad (A.5)$$

## References

- Anderson, P.M. & Burgess, S.M. (2000): Empirical Matching Functions: Estimation and Interpretation Using Disaggregate Data. *Review of Economics and Statistics*, 82, 93-102.
- Bleakley H. & Fuhrer, J.C. (1997): Shifts in the Beveridge Curve, Job Matching and Labour Market Dynamics. *New England Economists Review* 0:9, 3-19.
- Broersma L. & Van Ours, J. (1998): Job Searchers, Job Matches and the Elasticity of Matching. *Labour Economics* 6, 77-93.
- Burgess, S.M. (1993): A Model of Competition between Unemployed and Employed Job Searchers: An Application to the Unemployment Outflow Rate in Britain. *Economic Journal* 103, 1190-204.
- Burgess, S.M. & Turon H. (2003): Unemployment Equilibrium and On-the-job Search. IZA Discussion Paper No. 753.
- Coles, M.G. & Smith, E. (1996): Cross-Section Estimation of the Matching Function: Evidence from England and Wales. *Economica* 63, 589-98.
- Fahr, R. & Sunde, U. (2001): Disaggregate Matching Functions. IZA Discussion Paper No. 335.
- Hämäläinen H. (2003): Työvoiman rekrytointi toimipaikoissa vuonna (2002). [in Finnish]. Ministry of Labour, Helsinki.

Ibourk A.; Maillard, B.; Perelman, S.; Sneessens H.R. (2004): Aggregate Matching Efficiency: A Stochastic Production Frontier, France 1990-1994. *Empirica* 31, 1-25.

Kano, S. & Ohta, M. (2005): Estimating a Matching Function and Regional Matching Efficiencies. *Japan and the World Economy* 17, 25-41.

Laird, N.M. & Ware, J.H. (1982): Random-effects Models for Longitudinal Data. *Biometrics*, 38, 963-974.

Mumford, K. & Smith, P.N. (1999): The Hiring Function Reconsidered: on Closing the Circle. *Oxford Bulletin of Economics and Statistics* 61, 343-364

Pissarides, C.A. (1994): Search Unemployment with On-the-job Search. *Review of Economic Studies* 61, 457-475.

Räisänen H. (2004): Työvoiman hankinta julkisessa työnvälityksessä. [in Finnish]. Government Institute for Economic Research Reports No. 107.

StataCorp. (2001): Stata Statistical Software: release 7.0. College Station, TX: Stata Corporation.

Van Ours, J. (1995): An empirical note on employed and unemployed job search. *Economic Letters* 49, 447 – 472.

Verbege G. & Molenberghs G. (2001): Linear Mixed Models for Longitudinal Data. Springer Series in Statistics. Springer-Verlag New York, Inc.



Wahba, J. & Zenou, Y (2003): Density, Social Networks and Job Search Methods: Theory and Application to Egypt. CEPR Discussion Paper 3967, London.

Wall, H. J. & Zoega, G. (2002): The British Beveridge Curve: A Tale of Ten Regions. Oxford Bulletin of Economics and Statistics, 64, 261-80.