COMPOSITION OF THE JOB-SEEKER STOCK IN LABOUR MARKET MATCHING: A STOCHASTIC FRONTIER APPROACH[1]

Sanna-Mari Hynninen
University of Jyväskylä, School of Business and Economics
P.O. Box 35, FI-40014 University of Jyväskylä, Finland
email: sanna-mari.hynninen@econ.jyu.fi

Abstract: This paper investigates the technical efficiency of labour market matching taking a stochastic frontier approach. The data set consists of monthly data from 145 Local Labour Offices (LLOs) in Finland over the period 1995/01-2004/09. The true fixed-effects model is utilised in order to separate cross-sectional heterogeneity from inefficiency. According to the results, there are notable differences in matching efficiency between regions, and these differences contribute significantly to the number of filled vacancies. If all regions were as efficient as the most efficient one, the number of total matches per month would increase by over 10 %. If inefficiency had no role in the matching function, the number of matches would increase by almost 24 %. The weight of the composition of the job-seeker stock and other environmental variables in the determination of matching inefficiency is on average 61 %. In particular, job seekers out of the labour force and highly educated job seekers improve technical efficiency in the matching function.

Keywords: matching function, technical efficiency, heterogeneity, local labour markets

JEL: J24, J64

---

[1] This study was funded by the Academy of Finland and Yrjö Jahnsson Foundation. I am grateful to Aki Kangasharju, Jaakko Pehkonen, Ari Hyytinen and Mika Kortelainen for helpful comments and suggestions.

1. Introduction

Labour markets are commonly characterised by a large number of individuals searching for new jobs simultaneously with a large number of firms searching for new workers. This phenomenon is due to frictions in the matching process: job seekers and vacant jobs do not match immediately. To a certain extent, frictions are necessary to guarantee the quality of matches, but at worst they slow down the matching process yielding higher structural unemployment: job seekers do not match the available vacancies. Reasons behind the inefficiency of matching can be related to skill mismatch between job seekers and vacant jobs, to regional mismatch problems, to low search effort by job seekers, to ranking behaviour by firms, to impediments in the transmission of information, to wide heterogeneity of job seekers and firms in the labour market, and to inefficiency in the functioning of employment agency (e.g. Broersma and Van Ours 1999; Pissarides 1994; Anderson and Burgess 2000; Petrongolo and Pissarides 2001; Hynninen and Lahtonen 2007).

The qualitative matching of inputs is a crucial determinant of matching efficiency, as it determines whether or not a contact between a job seeker and a vacancy leads to a match. Therefore, in this study we focus on the role of the composition of the job-seeker stock in matching efficiency. We take a stochastic frontier approach to labour market matching in Finland (Coelli et al 1999; Kumbhakar and Lovell 2000). The concept of technical efficiency in the production function, presented in detail in Farrell (1957), is in the matching function determined by the ability of regions to produce matches by the stocks of job seekers and vacant jobs (Fahr and Sunde 2002; Ilmakunnas and Pesola 2003; Ibourk et al. 2004; Fahr and Sunde 2005). The matching function is interpreted as a frontier that determines the upper boundary for successful matches that could be produced by the given stocks of job seekers and vacant jobs.

The traditional fixed-effects model provides time-invariant estimates for efficiency relative to the best in the sample (Kim and Schmidt 2000). The problem in this approach is that all time-invariant heterogeneity across cross-sections is included in the efficiency term. Stochastic frontier analysis (SFA) avoids this problem of misspecification by providing a tool for the separation of efficiency from heterogeneity (Greene 2005a and b). Efficiency is also allowed to vary over time, which is a realistic assumption in long time series. In addition, a model specification of the Battese and Coelli (1995) type allows efficiency terms to be functions of variables that cause frictions in the matching process.

Ilmakunnas and Pesola (2003) and Hynninen et al. (2006)[2] have previously applied stochastic frontier analysis to the production of hires from unemployment in Finland. In this study, we investigate the efficiency of the production of filled vacancies. We apply Greene's (2005a and b) true fixed-effects stochastic frontier model with the inefficiency terms of the Battese and Coelli (1995) type. We utilise estimated efficiencies in order to calculate the quantitative effects of total inefficiency on matches. The matching function represents the production of filled vacancies during a month with job seekers and vacant jobs as inputs. The data are monthly panel data from 145 Local Labour Offices (LLOs) in Finland from the period 1995/01 – 2004/09[3]. The data consist of registered job seekers, vacant jobs and filled vacancies reported in state-run LLOs[4]. The data provide information on the composition of the job-seeker stock according to labour market status, age, and education.

The paper is organised as follows: Section 5.2 introduces the stochastic frontier approach to the matching function and specifies the models, Section 5.3 describes the data set, Section 5.4 discusses the results of the efficiency analysis and Section 5.5 concludes. Notable regional differences in efficiency were found. According to the results, aggregate level matches would increase by over 10 % if all regions were as efficient as the most

[2] Hynninen et al. (2006) studies the technical efficiency of hiring processes and the contribution of inefficiencies to the aggregate unemployment rate in 19 largest travel-to-work areas (TTWAs) in Finland. The study finds substantial efficiency differences between TTWAs, which further contribute significantly to the aggregate unemployment rate, i.e. 2.5 percentage points.

[3] Åland Island is excluded from the analysis due to its exceptional labour market conditions.

[4] The state-run employment agencies play an important role in the Finnish labour market. The proportion of jobs mediated by LLOs varied between a low of 49 in 1993 and a high of 71 per cent in 1996 over the period 1993-2002 (Hämäläinen 2003). The mean was around 60 per cent.

efficient one. If there were no inefficiency at all in the matching, the number of filled vacancies would increase by almost 24 %. In the job-seeker stock, job seekers out of the labour force and highly educated job seekers make the most important contribution to matching efficiency by notably increasing it.

2. Specification of the stochastic frontier matching model

We assume that labour market matching follows the production process determined by the familiar Cobb-Douglas production function (Pissarides 2000):

$$M_{i,t} = AS^{\alpha}{}_{i,t-1}V^{\beta}{}_{i,t-1}, \tag{1}$$

where $M_{i,t}$ denoted filled vacancies (vacancy outflow) during a month t in LLO i, $S_{i,t-1}$ the job-seeker stock and $V_{i,t-1}$ the stock of vacancies at the end of the previous month.

The stochastic logarithmic production frontier model takes the following form, defined by Battese and Coelli (1995) and Greene (2005a and b):

$$\ln M_{i,t} = [\mu_i + \alpha \ln S_{i,t-1} + \beta \ln V_{i,t-1}] + v_{i,t} - u_{i,t} \tag{2}$$

The expression in square brackets states the matching frontier that gives the maximum output, matches, which can be achieved at given amounts of production inputs, job seekers and vacancies. According to Greene (2005a and b) the model can be called the true fixed-effects model since it separates the true fixed effect μ_i from inefficiency $u_{i,t}$. In other words, time-invariant cross-sectional heterogeneity in the production of matches is separated from the inefficiency that causes deviations from the frontier. This decomposition is not possible in the basic fixed-effects models.

The observable error term $\varepsilon_{i,t} = v_{i,t} - u_{i,t}$ consists of two components that we do not directly observe. The "normal" error terms $v_{i,t}$ are iid and follow the $N(0, \sigma_v{}^2)$ distribution. $u_{i,t}$ are non-negative random variables accounting for technical inefficiency in the production of matches. They are assumed to be distributed independently of $v_{i,t}$, following the $N(Z_{j,it}\delta_j, \sigma_u{}^2)$ distribution truncated at zero (Coelli 1997). The $Z_{j,it}$ vector denotes inefficiency regressors and δ_j s are coefficients to be estimated. The variance of the composed error term is expressed as $\sigma^2 = \sigma^2{}_v + \sigma^2{}_u$. The relative importance of the residual associated to the inefficiency term is $\gamma = \sigma^2{}_u / \sigma^2$. σ^2 and γ are parameters to be estimated instead of $\sigma^2{}_v$ and $\sigma^2{}_u$.

The distribution of the inefficiency terms is effected by "environmental factors" that vary between cross-sectional units and over time. The inefficiency term is a function of these environmental factors, $u_{i,t} = Z_{j,it}\delta_j + w_{i,t}$, where the random variable $w_{i,t}$ is defined by the truncation of the normal distribution with zero mean and variance σ_u^2 such that the point of truncation is $-Z_{j,it}\delta_j$, i.e. $w_{i,t} \geq -Z_{j,it}\delta_j$. These assumptions are consistent with $v_{i,t}$ being non-negative truncations of the $N(Z_{j,it}\delta_j, \sigma_u{}^2)$ distribution (Battese and Coelli 1995). This specification assumes that all environmental factors that might increase or decrease inefficiency in the production influence directly the degree of technical efficiency, not the shape of the production technology as in the conventional fixed effects framework (Coelli et al. 1999).

The parameters of the stochastic frontier and the efficiency term can be estimated jointly by maximising the log-likelihood of the model (Coelli 1997; Coelli et al. 1998). The conditional estimates of the efficiency coefficients $TE_{i,t}$ are computed as

$$TE_{i,t} = [\exp(-u*_{i,t}) \mid M, S, V, Z]. \tag{3}$$

The efficiency measure is absolute, not relative to the best in the sample. It is equal to 1 when matches lie on the frontier, otherwise $TE_{i,t} < 1$.

3. Data description

The data comprise filled vacancies during a month and the stocks of registered job seekers and vacant jobs at the end of a previous month from 145 Local Labour Offices (LLOs) in Finland. The research period spans from January 1995 to September 2004. Following the examples of Fahr and Sunde (2002, 2005), Ilmakunnas and Pesola (2003), and Ibourk et al (2004) we include in the model control variables that capture labour market heterogeneity and possibly affect technical efficiency of the production of matches. These inefficiency regressors consist of the structure of the job-seeker stock according to labour market status, age, and education. Shares of long-term unemployed (over one year), job seekers out of the labour force, employed job seekers, job seekers below 25 years and over 50 years and primary educated as well as highly educated job seekers are included in the inefficiency terms.

Table 1 summarises the descriptive statistics of the data by LLOs. On average in a LLO there are 4 066 job seekers and 114 vacant jobs. A large share of the job-seeker pool, 14 % on average, is of long-term type. Employed job seekers account for 24 % and job seekers out of the labour force for 9 % of the job-seeker stock. By age, almost 20 % of job seekers are over 55 years old and 7 % are younger than 25 years. In educational composition the registered job seekers are predominantly the primary and secondary educated; only 9 % are highly educated (see Appendix 1 for the educational classification).

Table 1. Descriptive statistics

	Average	Min / Max	Std. Dev.
Matching rate	0.04	0 / 1.0	0.04
Filled vacancies	142	0 / 7 717	426
Job seekers	4 066	183 / 106 329	7 809
Vacant jobs	114	0 / 7 566	370
Share long-term unemployed	0.14	0.01 / 0.33	0.05
Share job seekers out of the labour force	0.09	0.01 / 0.44	0.05
Share employed job seekers	0.24	0.08 / 0.47	0.05
Share job seekers < 25 years	0.07	0 / 0.2	0.03
Share job seekers > 50 years	0.18	0.08 / 0.31	0.03
Share primary educated job seekers	0.49	0.32 / 0.7	0.06
Share highly educated job seekers	0.09	0.01 / 0.27	0.05

Figure 1. Matching probabilities and labour market tightness by LLOs

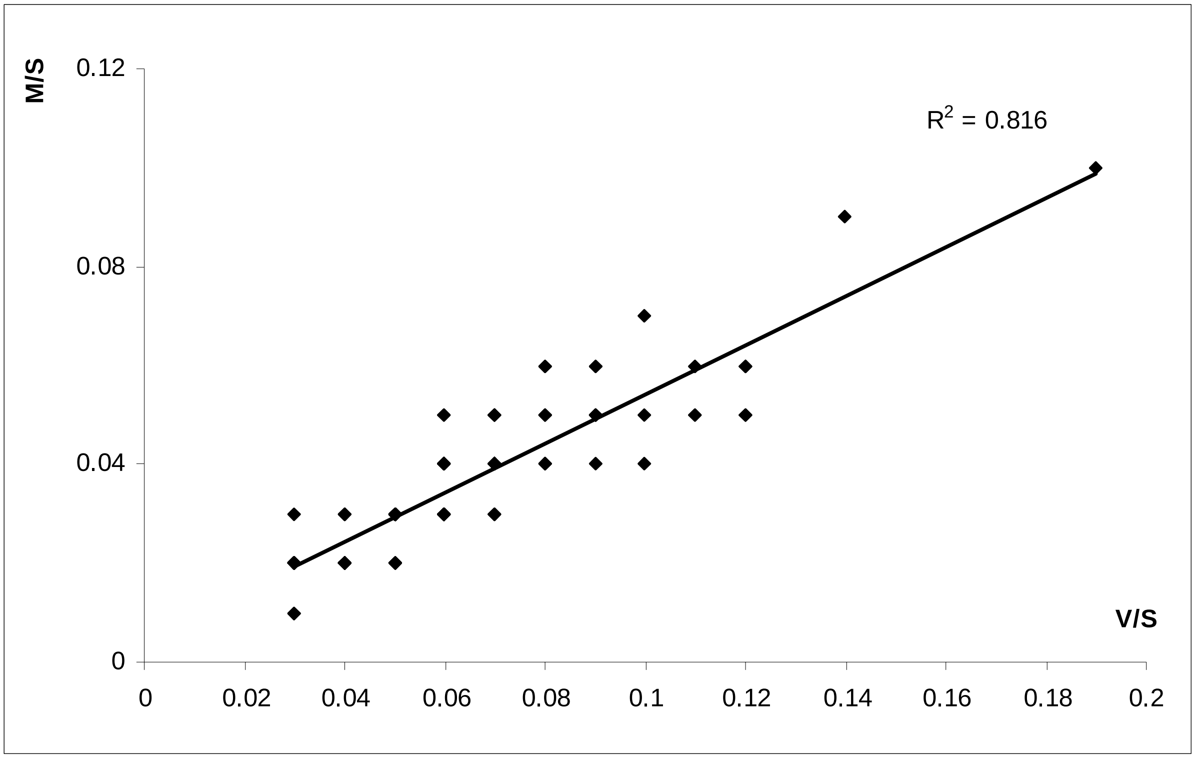
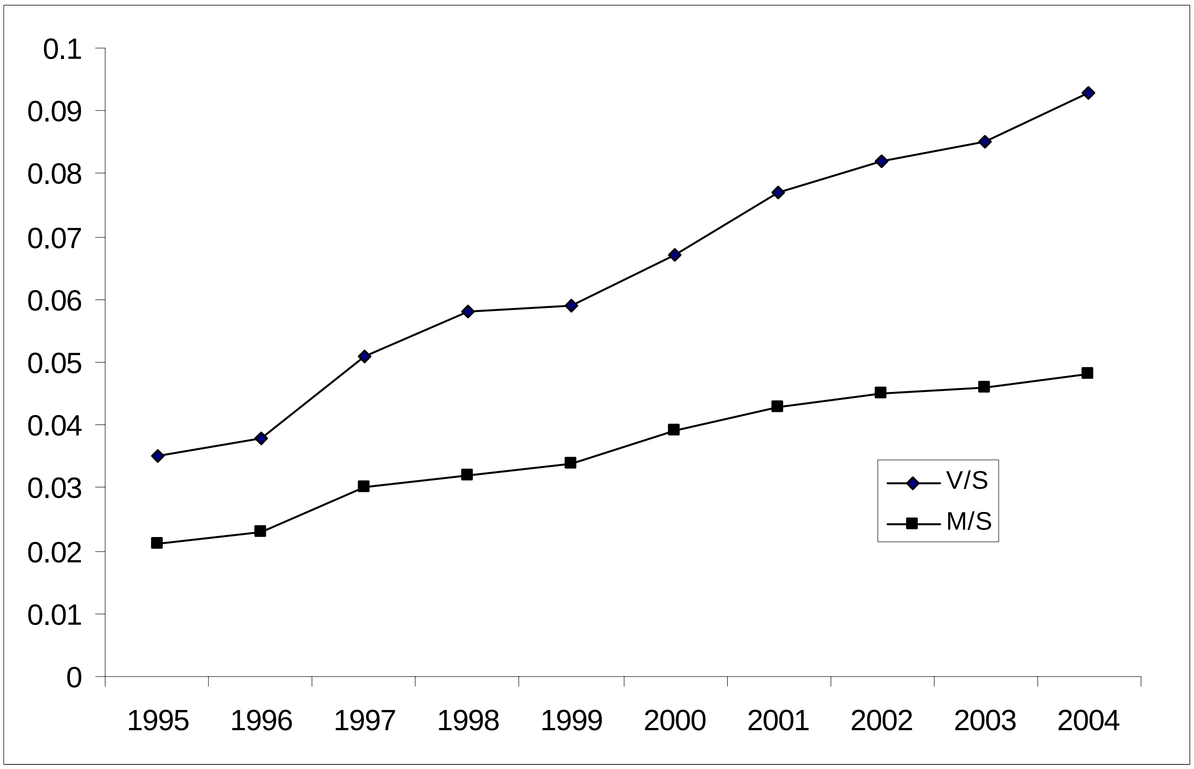


Figure 1 provides preliminary information on regional differences by tabulating matching probabilities (M/S) and labour market tightness (V/S) across regions[5]. The relationship between matching probability and tightness is clear: R^2 =0.82. The picture indicates differences in matching efficiency: at a given tightness LLOs produce deviating amounts of matches. Figure 2 in turn describes the changes in matching probability and labour market tightness by years. Both factors have increased continuously over the period. The change in the matching rate was notably slower, especially in the early 2000s. As a result, the gap between the matching rate and tightness also widened over the period. This indicates deterioration in matching efficiency: at a given labour market tightness the local labour markets are able to produce fewer matches. These figures furnish a starting point

[5] The flow of new vacant jobs during a month is included in the tightness in the figure. Owing to simultaneity bias problems, they are not, however, used in the matching function estimations. See Gregg and Petrongolo (2005) for stock-flow matching.

for our stochastic frontier analysis, which takes into account factors affecting efficiency and allows for time-wise variation in the efficiency estimates.

Figure 2. Matching probabilities and labour market tightness by years



4. Results

Five alternative specifications are reported in Table 2. Specification 1 is a conventional random-effects model and specification 2 a fixed-effects model. Specifications 3-5 are different kinds of stochastic frontier models. Specification 3 is a SFA model of the Battese and Coelli (1995) type without any panel-specific effects. Cross-sectional heterogeneity is added into the model through the inefficiency regressors, where it enters into the mean of the distribution of the inefficiency effects. Model 4 combines the Battese and Coelli -type of inefficiency effects with Greene's (2005a and b) true fixed-effects model by adding LLO-specific dummies into the function to capture time-invariant heterogeneity in the matching production. Specification 5, in addition, includes the time trend in the inefficiency term. In addition, in order to capture cyclical and seasonal variation in the matching function, we include yearly and monthly dummies in the function in all of the models[6].

Table 2. Estimation results

Variables	Conventional panel data models: Random	Conventional panel data models: Fixed	Stochastic frontier models: Battese and Coelli	Stochastic frontier models: True fixed 1	Stochastic frontier models: True fixed 2
Dependent variable: ln M_t	(1)	(2)	(3)	(4)	(5)
ln S_{t-1}	0.45***(0.02)	0.3***(0.06)	0.47***(0.01)	0.31***(0.05)	0.24***(0.05)
ln V_{t-1}	0.42***(0.005)	0.42***(0.005)	0.47***(0.005)	0.42***(0.005)	0.43***(0.005)
Constant	-1.75***(0.24)		-0.68		
Inefficiency controls			**-ln(efficiency)**		
t					0.025***(0.001)
(Share LTU)$_{t-1}$	-1.73***(0.2)	-1.21**(0.2)	4.68***(0.14)	0.97***(0.21)	-2.32***(0.45)
(Share OUT)$_{t-1}$	1.84***(0.19)	2.12***(0.2)	-0.41**(0.14)	-2.57***(0.21)	-9.28***(0.55)
(Share EMP)$_{t-1}$	0.3 (0.18)	0.47* (0.19)	1.01***(0.15)	-0.85***(0.2)	-4.44***(0.41)
(Share < 25)$_{t-1}$	0.48 (0.44)	0.64 (0.45)	-0.19 (0.35)	-1.7***(0.49)	-7.79***(1.01)
(Share > 50)$_{t-1}$	0.38 (0.32)	0.73*(0.35)	0.13 (0.24)	-0.1 (0.38)	-7.77*** (0.85)
(Share PRIMARY)$_{t-1}$	0.36 (0.21)	0.65** (0.23)	0.31**(0.12)	-0.31 (0.22)	4.13*** (047)
(Share HIGH)$_{t-1}$	2.64*** (0.3)	2.79*** (0.32)	-1.45***(0.2)	-3.83***(0.35)	-7.92***(0.35)
Constant				1.99***(0.22)	2.18***(0.35)
Returns to scale	0.87***	0.75***	0.94***	0.74***	0.67***
R^2	0.8	0.79			
Number of observations	16 965	16 965	16 965	16 965	16 965
sigma-squared			0.29	0.24	0.37
gamma			0.00006	0.3***	0.57**
log likelihood			-13 449	-11 479	-11 270
AIC			26 966	23 313	22 898
Hausman, Chi^2		146.5***			
LR-test, t=0, Chi^2					417***
Average efficiency			0.47 (0.12)	0.52 (0.16)	0.74 (0.17)

Notes: All models include yearly and monthly dummies in the function. Standard deviations reported in parentheses. *** 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. Of the conventional panel data models, Hausman test favours the fixed-effects model at the 0.1 % level. The LR test rejects the hypothesis that model 5 is nested in model 4 with a signifigance level of 0.1 %.

According to the results, the coefficient for vacancies is more stable across the specifications than the coefficient for job seekers, varying between 0.42 and 0.47. The job-seeker coefficient is more volatile, varying between 0.24 and 0.47. Random

[6] We also estimated all of the models with a trend in the function instead of yearly dummies. The models with dummies proved to have more explanatory power. The results on the estimations with a trend are available from the author.

specifications report notably higher job-seeker coefficients. They take into account between-units variation in addition to within-unit variation, which might yield the higher job-seeker coefficients. Among these conventional panel data models, the Hausman test, however, favours the fixed-effects specification against the random model. All models, independent of the type of panel effects or inclusion of the inefficiency terms, exhibit decreasing returns to scale.

The γ coefficients in the SFA models correspond to the estimated share of the inefficiency term in the variance of the composed error term, i.e., it is an indication of two-sided errors. In the Battese and Coelli specification the inefficiency term is insignificant, since γ is almost zero and not statistically significant. This indicates that all deviations from the frontier are due to random errors $v_{i,t}$ and that the model collapses to the basic OLS-model with inefficiency regressors in the matching function (Battese and Coelli 1995).

In the true fixed-effects model 1 (column 4) γ is 0.3 and highly significant, indicating that when we control for cross-sectional differences in the matching technology, stochastic inefficiency terms explain 30 % of the total variation in the composed error term[7]. This indicates that fixed effects are necessary in order to separate inefficiency effects; we have 145 cross-sections with wide heterogeneity. When, further, we add the time trend into the inefficiency term, γ rises to 0.57. Adding the time trend thus increases the fraction of inefficiency to the composed error term.

The log likelihood and AIC values favours specification 5 against the others. In addition, the likelihood ratio test rejects the hypothesis that the coefficient of the trend is zero. Hence, efficiency appears to have a negative trend, i.e. an exogenous decline occurred in matching efficiency during the period, as already indicated by the curves in Figure 2. It should be noted that adding the time trend has a marked affect on the results by

[7] The estimated inefficiency is clearly stochastic, not deterministic, which favours stochastic frontier analysis against data envelopment analysis, where all deviations from the frontier are assumed to be due to inefficiency.

decreasing the coefficient for job seekers in the function and attributing to the job-seeker stock variables in the inefficiency term more importance. This means that variations in the composition of the job-seeker stock contribute to efficiency notably more than in the model without the time trend.

4.1 Determinants of the matching efficiency

Many previous studies have reported that the search intensity of job seekers (e.g. Budd et al 1988; Layard and Bean 1989; Pissarides 1992) and the ranking behaviour of firms (Burgess 1993; Blanchard and Diamond 1994; Pissarides 1994; Van Ours 1995; Broersma 1997; Broersma and Van Ours 1999; Mumford and Smith 1999; Anderson and Burgess 2000; Burgess and Turon 2003) are crucial determinants of the size of the matching frictions. In line with this, we assume the matching inefficiency to be a linear function of the composition of the job-seeker stock. We control for the composition of the job-seeker stock regarding labour market position, age and education. With respect to labour market position, we define unemployed job seekers with an unemployment spell shorter than a year as the "base" group of job seekers with respect to age, job seekers aged between 26-49 years and with respect to education, secondary educated job seekers form the base groups. The efficiency effects of other groups are studied in relation to these base groups.

Our results for long-term unemployment are not straightforward. In a conventional fixed-effects model (Table 2, column 2) long-term unemployment negatively affects matches by the coefficient -1.21, as expected. According to the true fixed-effects model 1 (Specification 4), a one percentage point increase in the group of long-term unemployed decreases matching efficiency by about 1 %[8]. Adding the time trend into the true fixed-effects model, however, changes the sign and magnitude of long-term unemployment (Specification 5): according to that specification, a one percentage point increase in long-term unemployment increases matching efficiency by over 2 %. Evidently, the negative

[8] Note that in SFA models a negative sign means a positive effect on efficiency: inefficiency = -(ln efficiency)

time trend captures the efficiency-decreasing effect of an increase in long-term unemployment. Long-term unemployment fell continuously during the research period, while efficiency also fell: the correlation between the trend and long-term unemployment is -0.40. The result is in line with Ilmakunnas and Pesola (2003) who report that long-term unemployment has a positive effect on hiring efficiency in Finland. Either Blanchard and Diamond (1989) did not find a statistically significant negative effect of long-term unemployment on matches.

The unequal employability of different job-seeker groups is clearly implied by the results for job seekers out of the labour force and employed job seekers. Both of these groups reduce matching frictions in LLOs. The negative inefficiency effect of job seekers out of the labour force is over two times larger than that of employed job seekers (Specification 5). This reflects that job seekers trying to enter the labour market are favoured by employers possibly due to their flexibility and freshness of skills that at least lately graduated students have. Their own search effort might also be higher than the effort of other groups. The same explanations hold for the efficiency enhancing effect of young job seekers who have found to improve efficiency also in Fahr and Sunde (2002) in Western Germany and in Ilmakunnas and Pesola (2003) in Finland. Older job seekers also improve efficiency in LLOs reflecting the value accorded the experience of older job seekers by firms seeking workers through state-run employment agencies.

The educational structure of the job-seeker stock is also of significance. The share variables capture the effect of primary and highly educated job seekers in relation to the secondary educated. A one percentage point increase in the high education group increases efficiency by almost 8 %. This is in line with results of Lahtonen (2006) in Finland and with those of Fahr and Sunde (2002) in the SFA framework in Western Germany. Fahr and Sunde argue that highly educated job seekers might have a higher search intensity and that the search process may be more directed in the high-education segment of the labour market, thereby contributing to higher matching efficiency.

Primary educated job seekers seem to decrease matching efficiency. These results could indicate job competition between job seekers with different levels of education. Employers might prefer highly educated to primary and secondary educated job seekers even where the job does not necessarily require high education. The existing evidence on job competition is not, however, unproblematic (Sicherman 1991; Van Ours and Ridder 1995; Gautier et al. 2002): Van Ours and Ridder found evidence of job competition between academic and higher vocational education, but not at lower levels of education, while others found no educational-related evidence of job competition.

4.2 Quantitative effects of inefficiency on matches

The average efficiency levels vary from 0.47 in a Battese and Coelli to 0.74 in the true fixed effects model with the time trend in the inefficiency term. (Table 2). We face the familiar problem that the efficiency estimates are not robust across SFA models, as previously reported, e.g., in Giannakas et al. (2003). Both the LR test for the significance of the trend in the inefficiency estimates and the AIC favour specification 5, as already reported above. On the basis of these tests we end up using the estimates given by them in our further calculations.

Regional variation in the mean efficiency varies from 0.36 to 0.89 (Appendix 2). If we consider all 16 965 efficiency estimates, the variation ranges from 0.06 to 0.95 with a standard deviation 0.17. On average, the matching process works rather efficiently; however, there are also inefficient regions which are permanently far from the frontier. The ranking of regions according to efficiency remained, however, rather stable during the research period: the Spearman rank correlation coefficient between the estimates for 1995 (the first year) and 2003 (last full year of the period) is 0.74.

We clarify the quantitative dimension of regional inefficiency from a somewhat different perspective from that of Ibourk et al. (2004) who also calculate efficiency slacks and the explanatory power of environmental variables. Our focus is on the magnitude of inefficiency and its direct effects on the number of monthly matches. Table 3 reports the results of those calculations. If there was no inefficiency at all, i.e., the efficiency level were 1 in all regions, we would obtain 2 727 more filled vacancies in a month. This implies a 23.7 % monthly increase in matches compared to the level of matches obtained at the current average levels of inefficiency. Comparing the number of matches obtained at the prevailing inefficiencies with the hypothetical number of matches obtained with zero-level inefficiency implies that inefficiency decreases matches by 19.2 %.

It is, however, unrealistic to assume that inefficiency plays a zero-role in the matching function. It is more appropriate to set the efficiency frontier at the highest level found in the sample. The highest average efficiency level, 0.89, is obtained in Vaasa (in Ostrobothnia). If we set all LLOs at the efficiency level of Vaasa we would achieve 1 174 new matches in a month, which would increase matches by 10.2 %. Comparing the number of matches obtained at the prevailing inefficiencies with the hypothetical number of matches obtained with Vaasa's inefficiency implies that inefficiency decreases matches by 9.3 %.

Table 3. Quantitative effects of inefficiency on matches

Efficiency calculations		Increase in matches
Predicted matches	11 484	
Matches with highest efficiency in the sample	12 658	1 174, 10.2 %
Matches with efficiency level of 1	14 211	2 727, 23.7 %
The weight of Z-variables in the inefficiency determination, mean	61 %	
Correlation between the weight of Z-variables and inefficiency level	0.25	

As defined in Section 2, the inefficiency estimates consist of two parts: $u_{i,t} = Z_{j,it}\delta_j + w_{i,t}$, i.e., of the part explained by inefficiency regressors and a random error. The Z variables contain the variables describing the composition of the job-seeker stock, the time trend and a constant. We have calculated the weight of the Z variables in the determination of the inefficiency estimates by comparing the absolute value of the inefficiency level predicted by the Z variables to the sum of this prediction and the absolute value of inefficiency predicted by random terms $w_{i,t}$ [9] (Appendix 2). The greater the particular absolute value, the greater the importance in the inefficiency term. According to the calculations, the weight of the $Z_{j,it}\delta_j$ set is on average 61 % in the inefficiency estimates. There is, however, weak positive dependence between the importance of the Z variables and the level of inefficiency: the correlation coefficient between inefficiency and the weight of the Z variables is 0.25. This indicates that, at lower levels of efficiency, the Z variables play a more important role while factors not related to the composition of the job-seeker stock become relatively more important at higher efficiency levels.

5. Conclusions

We studied the process of matching job seekers and vacant jobs in local labour markets taking a stochastic frontier approach. We applied true fixed-effects modelling in order to decompose the time-invariant cross-sectional heterogeneity that directly affects the matching technology from inefficiency that causes deviations from the frontier. The inefficiency terms were modelled as functions of the job-seeker stock composition in the regions.

Notable differences in matching efficiency between regions were found, and these differences were shown to have significant effects on the number of filled vacancies. If all regions were as efficient as the most efficient one, the number of total matches in a month would increase by over 10 %. If there were no inefficiency at all in the matching function, matches would increase by almost 24 %.

The results indicate that a continuous exogenous decline in matching efficiency occurred during the research period. The results also show that changes in the composition of the

[9] Note that parts of the efficiency estimates can predict negative inefficiency. Together they determine level of inefficiency higher than 0 (Battese and Coelli 1995).

job-seeker stock strongly contribute to the efficiency estimates: the labour market status, age as well as educational structure of the job-seeker stock strongly affect the ability of local labour markets to form successful matches. In particular, job seekers out of the labour force and highly educated job seekers improve matching efficiency. The total weight of the set of inefficiency regressors in the inefficiency term is on average 61 %.

References

Anderson, P. and Burgess, S. (2000) Empirical matching functions: Estimation and interpretation using state-level data. *The Review of Economics and Statistics* 82, 93-102.

Battese, G.E. and Coelli, T.J. (1995) A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics* 20, 325-332.

Blanchard, O. and Diamond, P. (1989) The Beveridge curve. *Brookings Papers on Economic Activity* 1, 1-76.

Blanchard O. and Diamond P. (1994) Ranking, unemployment duration, and wages. *Review of Economic Studies* 61, 417-434.

Broersma L. (1997) Competition between employed and unemployed job searchers: is there a difference between the UK and The Netherlands? *Applied Economics Letters* 4, 199-203.

Broersma, L. and Van Ours, J. (1999) Job searchers, job matches and the elasticity of matching. *Labour Economics* 6, 77-93.

Budd A., Levine P., and Smith P. (1988) Unemployment, vacancies, and the long-term unemployed. *Economic Journal* 98, 1071-1091.

Burgess S. (1993) A model of competition between unemployed and employed job searchers: an application to the unemployment outflow rate in Britain. *Economic Journal* 103, 1190-1204.

Burgess S. and Turon H. (2003) Unemployment equilibrium and on-the-job search. *IZA Discussion Paper* No. 753.

Coelli, T. (1997) A guide to FRONTIER Version 4.1: A Computer program for stochastic frontier production and cost estimation. *CEPA Working paper* 96/07.

Coelli, T., Perelman, S., and Romano, E. (1999) Accounting for environmental influences in stochastic frontier models: with application to international airlines. *Journal of Productivity Analysis* 11, 251-273.

Coelli, T., Rao D.S., and Battese G. (1998) *An introduction to efficiency and productivity analysis*. Kluwer Academic Publishers.

Fahr, R. and Sunde, U. (2002) Estimations of occupational and regional matching efficiencies using stochastic production frontier models. *IZA Discussion paper* 552.

Fahr, R. and Sunde, U. (2005) Regional dependencies in job creation: an efficiency analysis for western Germany. *IZA Discussion paper* 1660.

Farrell, M.J. (1957) The measurement of productive efficiency. *Journal of the Royal Statistical Society Series A* 120(3), 253-290.

Gautier, P., van den Berg, G., Van Ours, J., and Ridder, G. (2002) Worker turnover at the firm level and crowding out of lower educated workers. *European Economic Review* 46, 523-538.

Giannakas, K., Tran, K., and Tzouvelekas, V. (2003) On the choice of functional form in stochastic frontier modelling. *Empirical Economics* 28, 75-100.

Greene, W. (2005a) Fixed and random effects in stochastic frontier models. *Journal of Productivity Analysis* 23, 7-32.

Greene, W. (2005b) Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics* 126, 269-303.

Gregg and Petrongolo (2005) Stock-flow matching and the performance of the labor market. *European Economic Review* 49, 1987-2011.

Hynninen, S., Kangasharju, A., and Pehkonen, J. (2006). Regional Matching Frictions and Aggregate Unemployment. *VATT Discussion Paper* 383.

Hynninen, S. and Lahtonen, J. (2007) Does population density matter in the process of matching heterogeneous job seekers and vacancies? Forthcoming in *Empirica.*

Hämäläinen H. (2003) Työvoiman rekrytointi toimipaikoissa vuonna 2002, [in Finnish], *Ministry of Labour*, Helsinki.

Ibourk, A., Maillard, B., Perelman, S., and Sneessens, H. (2004) Aggregate matching efficiency: A stochastic production frontier approach, France 1990-1994. *Empirica* 31, 1-25.

Ilmakunnas, P. and Pesola, H. (2003) Regional labour market matching functions and efficiency analysis. *Labour* 17 (3), 413-437.

Kim, Y. and Schmidt, P. (2000) A review and empirical comparison of Bayesian and classical approaches to inference on efficiency levels in stochastic frontier models with panel data. *Journal of Productivity Analysis* 14, 91-118.

Kumbhakar, S. and Lovell, C.A. (2000) *Stochastic frontier analysis*. Cambridge University press.

Lahtonen, J. (2006) *Matching heterogeneous job seekers and vacancies: Empirical studies using Finnish data*. Jyväskylä Studies in Business and Economics 50.

Layard R. and Bean C. (1989) Why does unemployment persist? *Scandinavian Journal of Economics* 91, 371-396.

Mumford K. And Smith P. (1999) The hiring function reconsidered: on closing the circle. *Oxford Bulletin of Economics and Statistics* 61, 343-364.

Petrongolo B. and Pissarides C. (2001) Looking into the black box: a survey of the matching function. *Journal of Economic Literature* 39, 390-431.

Pissarides C. (1992) Loss of skill during unemployment and the persistence of employment shocks. *Quarterly Journal of Economics* 107, 1371-1391.

Pissarides C. (1994) Search unemployment with on-the-job search. *Review of Economic Studies* 61, 457-475.

Pissarides C. (2000) *Equilibrium unemployment theory*. MIT Press.

Sicherman, N. (1991) “Overeducation” in the labor market. *Journal of Labor Economics* 9, 101-122.

Van Ours J. (1995) An empirical note on employed and unemployed job search. *Economics Letters* 49, 447-452.

Van Ours, J. and Ridder, G. (1995) Job matching and job competition: Are lower educated workers at the back of queues? *European Economic Review* 39, 1717-1731.

Appendix 1. Relation between 3-group classification and ISDEC 1997

ISDEC 1997	Name	3-group classification
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

Appendix 2. Efficiency and inefficiency by LLOs

LLO	Efficiency	Inefficiency	Ineff. due Z-variables	Ineff. due other factors	Weight of Z-variables
Average	**0.737**	**0.324**	**0.792**	**-0.468**	**0.607**
Vaasa	0.891	0.116	-0.242	0.358	0.404
Tuusula	0.884	0.123	0.054	0.069	0.436
Jyväskylä	0.880	0.127	-0.292	0.419	0.410
Parainen	0.880	0.128	0.020	0.107	0.159
Kaarina	0.876	0.133	0.290	-0.157	0.649
Eura	0.875	0.134	0.746	-0.612	0.549
Tampere	0.873	0.136	-0.078	0.214	0.268
Helsinki	0.872	0.137	-0.083	0.220	0.274
Turku	0.870	0.139	-0.023	0.163	0.126
Ylöjärvi	0.862	0.148	0.302	-0.154	0.662
Kuopio	0.862	0.148	-0.161	0.309	0.343
Naantali	0.861	0.150	0.096	0.053	0.644
Valkeakoski	0.860	0.151	0.343	-0.192	0.641
Raisio	0.858	0.153	0.307	-0.154	0.666
Seinäjoki	0.857	0.154	-0.168	0.322	0.343
Hämeenlinna	0.857	0.154	0.257	-0.103	0.714
Mikkeli	0.857	0.155	0.090	0.065	0.579
Pietarsaari	0.856	0.155	0.466	-0.310	0.600
Oulu	0.856	0.156	-0.164	0.320	0.339
Kotka	0.855	0.157	0.128	0.029	0.818
Rauma	0.853	0.159	0.389	-0.230	0.628
Kouvola	0.853	0.159	0.318	-0.158	0.667
Pori	0.851	0.161	0.360	-0.198	0.645
Vihti	0.851	0.161	0.127	0.034	0.788
Varkaus	0.850	0.162	0.243	-0.080	0.752
Lahti	0.849	0.163	0.506	-0.343	0.596
Kirkkonummi	0.849	0.163	0.288	-0.124	0.698
Lappeenranta	0.849	0.164	0.360	-0.196	0.647
Mänttä	0.847	0.166	0.731	-0.565	0.564
Hyvinkää	0.846	0.167	0.426	-0.260	0.622
Paimio	0.846	0.167	0.169	-0.002	0.990
Kerava	0.845	0.168	0.554	-0.386	0.589
Kokkola	0.842	0.172	0.243	-0.071	0.774
Kajaani	0.841	0.173	0.078	0.095	0.450
Kemiö	0.841	0.173	0.860	-0.687	0.556
Kangasala	0.841	0.173	0.669	-0.496	0.574
Karjaa	0.840	0.174	0.512	-0.338	0.603
Salo	0.839	0.175	0.651	-0.476	0.578
Rovaniemi	0.838	0.177	0.196	-0.019	0.912
Loimaa	0.835	0.180	0.601	-0.422	0.588
Jämsä	0.835	0.181	0.734	-0.553	0.570
Kauhava	0.834	0.181	0.311	-0.130	0.706
Hamina	0.834	0.181	0.371	-0.190	0.662
Porvoo	0.834	0.182	0.400	-0.218	0.647
Lempäälä	0.832	0.184	0.555	-0.372	0.599
Joensuu	0.831	0.185	0.026	0.159	0.138
Kuusankoski	0.830	0.187	0.985	-0.799	0.552

LLO	Efficiency	Inefficiency	Ineff. due Z-variables	Ineff. due other factors	Weight of Z-variables
Lohja	0.828	0.188	0.696	-0.507	0.578
Lapua	0.826	0.191	0.511	-0.320	0.615
Kristiinankaupunki	0.823	0.195	0.543	-0.349	0.609
Harjavalta	0.823	0.195	0.668	-0.472	0.586
Järvenpää	0.821	0.197	0.378	-0.182	0.676
Tammisaari	0.817	0.202	0.637	-0.436	0.594
Vammala	0.817	0.202	0.464	-0.262	0.639
Forssa	0.817	0.203	0.961	-0.758	0.559
Janakkala	0.816	0.204	0.660	-0.456	0.591
Ääneseutu	0.812	0.208	0.779	-0.571	0.577
Nurmijärvi	0.809	0.211	0.772	-0.560	0.579
Savonlinna	0.808	0.213	0.457	-0.244	0.652
Kemi	0.808	0.214	0.513	-0.299	0.632
Imatra	0.807	0.214	1.113	-0.899	0.553
Heinola	0.806	0.215	0.792	-0.576	0.579
Riihimäki	0.805	0.217	0.757	-0.541	0.584
Kurikka	0.804	0.218	0.572	-0.354	0.618
Lappajärvi	0.803	0.220	0.747	-0.527	0.586
Hämeenkyrö	0.796	0.228	0.565	-0.337	0.626
Nokia	0.796	0.228	0.737	-0.508	0.592
Anjalankoski	0.794	0.231	0.780	-0.549	0.587
Uusikaupunki	0.793	0.232	0.896	-0.663	0.575
Pieksämäki	0.791	0.235	0.482	-0.247	0.661
Leppävirta	0.790	0.236	0.446	-0.210	0.680
Laitila	0.789	0.237	1.053	-0.815	0.564
Kyrönmaa	0.785	0.242	0.479	-0.237	0.669
Loviisa	0.785	0.243	1.002	-0.760	0.569
Mäntsälä	0.784	0.243	0.789	-0.546	0.591
Karkkila	0.784	0.243	0.930	-0.687	0.575
Siilinjärvi	0.783	0.245	0.354	-0.109	0.764
Ylivieska	0.779	0.250	0.495	-0.245	0.669
Huittinen	0.776	0.254	0.934	-0.680	0.579
Keuruu	0.775	0.255	0.917	-0.663	0.581
Raahe	0.769	0.263	0.399	-0.136	0.746
Alavus	0.764	0.269	0.754	-0.485	0.608
Jalasjärvi	0.763	0.271	0.588	-0.318	0.649
Haukipudas	0.754	0.282	0.649	-0.366	0.639
Orivesi	0.754	0.283	1.083	-0.800	0.575
Iisalmi	0.753	0.284	0.533	-0.250	0.681
Juva	0.752	0.285	0.886	-0.601	0.596
Laukaa	0.749	0.289	0.749	-0.460	0.620
Kaustinen	0.745	0.294	0.443	-0.149	0.748
Parkano	0.743	0.297	1.086	-0.788	0.579
Sotkamo	0.741	0.300	0.564	-0.264	0.681
Virrat	0.740	0.301	0.757	-0.455	0.624
Mäntyharju	0.738	0.304	1.403	-1.099	0.561
Toijala	0.736	0.307	1.140	-0.833	0.578
Haapavesi	0.726	0.321	0.701	-0.381	0.648
Saarijärvi	0.725	0.322	1.140	-0.817	0.582

LLO	Efficiency	Inefficiency	Ineff. due Z-variables	Ineff. due other factors	Weight of Z-variables
Kemijärvi	0.724	0.323	1.066	-0.743	0.589
Parikkala	0.721	0.327	1.186	-0.859	0.580
Noormarkku	0.719	0.331	1.236	-0.905	0.577
Tornio	0.712	0.339	0.814	-0.475	0.632
Kerimäki	0.707	0.346	0.931	-0.585	0.614
Ylitornio	0.705	0.350	1.089	-0.739	0.596
Suupohja	0.701	0.356	1.053	-0.698	0.602
Kuusamo	0.691	0.370	0.694	-0.324	0.682
Alajärvi	0.683	0.381	1.094	-0.713	0.605
Haapajärvi	0.683	0.382	0.923	-0.541	0.630
Muonio	0.682	0.382	1.104	-0.722	0.605
Hanko	0.682	0.383	1.072	-0.689	0.609
Heinävesi	0.669	0.402	1.200	-0.798	0.601
Kankaanpää	0.658	0.419	0.980	-0.561	0.636
Kuhmo	0.657	0.420	1.269	-0.849	0.599
Kangasniemi	0.654	0.425	1.095	-0.670	0.620
Kiuruvesi	0.632	0.458	0.835	-0.377	0.689
Utsjoki	0.621	0.476	0.973	-0.497	0.662
Lieksa	0.618	0.481	1.274	-0.793	0.616
Lapinlahti	0.616	0.485	1.275	-0.790	0.617
Keski-Karjala	0.615	0.487	1.277	-0.791	0.618
Viitasaari	0.610	0.494	1.294	-0.801	0.618
Tervola	0.602	0.507	1.639	-1.132	0.591
Vaala	0.579	0.546	1.366	-0.820	0.625
Outokumpu	0.578	0.547	1.230	-0.683	0.643
Nilsiä	0.565	0.571	1.446	-0.874	0.623
Joutsa	0.564	0.572	1.295	-0.722	0.642
Paltamo	0.562	0.576	1.293	-0.717	0.643
Ylä-Karjala	0.561	0.578	1.414	-0.836	0.628
Sisä-Savo	0.559	0.582	1.201	-0.619	0.660
Suomussalmi	0.557	0.585	1.498	-0.913	0.621
Kittilä	0.545	0.606	1.781	-1.175	0.603
Pielavesi	0.543	0.611	1.651	-1.040	0.614
Savukoski	0.542	0.612	1.445	-0.833	0.634
Eno	0.525	0.645	1.700	-1.055	0.617
Pello	0.521	0.652	1.731	-1.079	0.616
Hyrynsalmi	0.508	0.676	1.723	-1.046	0.622
Sodankylä	0.501	0.691	1.161	-0.470	0.712
Juankoski	0.493	0.708	1.500	-0.792	0.655
Kolari	0.486	0.722	1.730	-1.008	0.632
Posio	0.471	0.752	1.634	-0.882	0.650
Pudasjärvi	0.469	0.758	1.587	-0.829	0.657
Ivalo	0.455	0.787	1.322	-0.536	0.712
Ranua	0.431	0.841	1.514	-0.673	0.692
Karstula	0.429	0.845	1.417	-0.571	0.713
Salla	0.418	0.873	1.899	-1.026	0.649
Ilomantsi	0.412	0.887	1.955	-1.069	0.647
Puolanka	0.381	0.966	1.792	-0.826	0.685
Enontekiö	0.359	1.024	1.855	-0.831	0.691