Aggregate Matching Efficiency:

A Stochastic Production Frontier Approach, France 1990-1994

Aomar Ibourk[a], Bénédicte Maillard[b], Sergio Perelman[c] and Henri R. Sneessens[d*]

Revised version

November 2001

Abstract

We evaluate the determinants of aggregate matching efficiency changes through a *stochastic Cobb-Douglas production frontier* model. The efficiency coefficient is represented by a stochastic function of variables meant to capture workers and firms characteristics. The model is estimated on French data covering twenty-two regions from March 1990 till February 1995. Our estimates suggest that aggregate matching efficiency has decreased steadily in the early nineties. There are also wide cross-regional differences. On average, about 30% of the variations of efficiency observed across time and regions can be related to changes in the explanatory variables used in the model. The most important explanatory variables are the proportion of youngsters, females and immigrants in the stock of job seekers. Long-term unemployment has a significant negative effect, population density a significant positive one. The huge decline in the proportion of permanent job offers has apparently little effect on matching efficiency.

Keywords: matching efficiency, regional unemployment, stochastic frontier

JEL classification: J64, C24

[*a] CREQ, Université de Marrakech; [b,d] FLSE, Université catholique de Lille; [c] CREPP, Université de Liège; [d] IRES, Université catholique de Louvain, and IZA. We are grateful to Bruno Van der Linden and participants at the EEA and EALE meetings for useful comments and suggestions. This work has benefited from the financial support of the Belgian government (Pôles d'attraction inter-universitaires PAI IV) and the Communauté Française de Belgique (PAC 98/03-221).

1 Introduction

Since the mid-seventies, unemployment rates have remained at abnormally high levels in many EU countries. Typically, the rise in unemployment coincided with a deterioration in the relationship between unemployment and vacancies (an outward shift of the so-called Beveridge curve). The magnitude of these increases and shifts varies substantially across countries, and even more so across regions (see EC (1999), part I, section IV).

Beveridge curve shifts can be analysed and interpreted in terms of search behaviours and the matching of jobs and workers. The *matching function* concept provides a most parsimonious and powerful way to model labour market frictions and analyse their macroeconomic consequences (see Petrongolo-Pissarides (2001) for a survey of the theoretical and empirical issues related to the use of this concept). The intuition behind the matching function is that the matching process can be compared to a production process (typically a Cobb-Douglas function) whose output would be the number of matches (the flow of hirings) and the two inputs the number of job seekers and the number of vacancies. The efficiency of the matching process (total factor productivity in standard production analysis) determines the number of matches that will be observed at given input values. When the flow of matches compensates the flow of separations (quits and layoffs) and labour force growth, the aggregate unemployment and vacancy rates remain unchanged from one period to the next. In such a stationary situation, the matching function can be recast as a Beveridge curve.

A Beveridge curve shift will in this setup reflect either a declining matching efficiency or larger separation and labour force growth rates implying larger reallocation and matching efforts. It is thus interesting to try and evaluate whether changes in the matching efficiency may contribute to explain Beveridge curve differences observed over time and across regions. Several authors have examined this question by estimating aggregate or disaggregated (across regions, sectors or occupations) matching functions (see for instance Blanchard and Diamond (1989, 1990) and Anderson and Burgess (2000) for the US economy, Coles and Smith (1996) for England, Entorf (1998, chap.3) for Germany, Gorter and van Ours (1994) for the Netherlands, Maillard (1997) and Agullo (1999) for France, etc...). Our paper addresses the very same issue and uses to this end a stochastic production frontier approach. Because the matching function is usually interpreted as a production function and because the key issue is to estimate efficiency, the stochastic production frontier approach seems a most natural modelling strategy. Warren (1991) used this approach to estimate the frictional unemployment rate in US manufacturing. More recently, Ibourk and Perelman (2000) used a more elaborated frontier model proposed by Battese and Coelli (1995) to identify the factors that may explain matching efficiency differences across regional labour markets in Morocco. Our objective is to apply the same methodology on French panel data, covering twenty-two regions over the period 1990:3 till 1995:2. The method provides an estimate of the matching efficiency changes observed in both the time and space dimensions. It also evaluates what part of these changes can be attributed to variables like long-term unemployment, age structure, etc... This empirical exercise can be regarded as an attempt at providing further insights into the "matching function black box" (see Petrongolo-Pissarides (2001)).

The contribution of matching efficiency changes to observed unemployment rate changes remains a debated issue. Blanchard and Diamond (1989, 1990) estimate a matching function on aggregate US data over the 1968:2-1981:12 period. They obtain a negative trend effect. The latter cannot be explained by long term unemployment. Their conclusion is that efficiency changes contribute to explain unemployment changes at low frequencies; at high frequencies, aggregate activity shocks rather than efficiency changes (reallocation shocks) dominate the movement of unemployment. van Ours (1991) estimates a matching function on Dutch data over the period 1961-1987. The estimated matching efficiency is negatively related to changes in the replacement ratio and long term unemployment. It remains quite stable over time, except for an unexplained decrease in the late sixties. Gorter and van Ours (1994) examine regional differences across the Dutch economy in the eighties. They obtain substantial regional differences (estimated regional efficiency levels vary by a factor of two). These differences seem to be barely significant though, and not related to long term unemployment or occupational mismatch. M.Maillard (1997) obtains similar results for France over the period 1974-1994. Over a smaller period (1990:1-1994:12) and with a richer data set, Agullo (1999) obtains regional differences that can be related to structural variables like long term unemployment, skill mismatch, proportion of old workers or permanent contracts, etc... The main difference between her paper and ours is methodological. The use of the stochastic frontier approach allows a more detailed analysis of the determinants of regional matching efficiencies.

We start in Section 2 with a brief presentation of the model and the estimation technique. The data are presented in section 3, the empirical results in section 4. Our main conclusions are gathered in section 5.

2 The Model

Let the matching process be represented by the following function:

$$H_{it} = F(V_{it-1}, U_{it-1}) \, . \, e_{it} \qquad \text{with } F_u, F_v > 0 \, . \tag{1}$$

H_{it} denotes the total flow of hirings observed in region i during period t. Function F has the same interpretation as a production function. The flow of hirings is represented by a positive function of the initial total stocks of job vacancies V_{it-1} and of job seekers U_{it-1}. The number of hirings also depends on an efficiency parameter e_{it}, assumed to enter multiplicatively (see Cobb-Douglas case below) and allowed to vary both through time and space. With constant returns to scale, equation (1) can be recast as follows:

$$\frac{H_{it}}{U_{it-1}} = f(\frac{V_{it-1}}{U_{it-1}}) \, . \, e_{it} \, , \tag{2}$$

i.e. the proportion of unemployed workers hired per unit of time is a positive function of labour market tensions, measured by the ratio of vacancies and unemployment, scaled by the efficiency parameter e_{it}. By definition, net employment growth is equal to the hiring rate $h_{it} \equiv H_{it}/N_{it-1}$ minus the separation rate s, that is:

$$\frac{\Delta N_{it}}{N_{it-1}} = h_{it} - s \, . \tag{3}$$

If function $F(.)$ displays constant returns to scale and net employment growth is equal to labour force growth g (that is, the hiring rate is equal to the sum of the separation and the labour force growth rates: $h_{it} = s + g$), equation (1) can be recast as an inverse relationship between the unemployment and the vacancy rates, the so-called Beveridge curve. In this setup, the position of the Beveridge curve would depend on s, g and the value of the efficiency parameter e.

Our main objective in this paper is to estimate and explain the efficiency changes that may have taken place both over time and across regions. To the extent that the matching process is compared to a production process and because the emphasis is on matching efficiency estimates, specifying the empirical model as a stochastic production frontier model seems a most natural modelling strategy. The frontier approach has been used for a long time in theoretical analyses of firms' behaviour. A firm's technical efficiency is in this framework measured by the distance between the observed input-output combinations and those given by the production frontier, defined as the set of all efficient input combinations (Shephard (1953)). These concepts are illustrated in figure 1, for a given and fixed level of output. The downward-sloping curve corresponds to an isoquant: it determines for every value of input U the minimum amount of input V needed to obtain the desired output level H. Points B and C represent two such fully efficient input combinations. At unchanged output level, point A represents an inefficient input combination. Debreu (1951) and Farrell (1957) define the *technical efficiency* of a firm operating at A as one minus the maximum equiproportional input reduction compatible with an unchanged output, *i.e.*, the ratio OB/OA. Our objective is to use this conceptual approach to evaluate how far observed matching outcomes may be from the efficiency frontier, and simultaneously to test what factors may contribute to explain these inefficiencies.

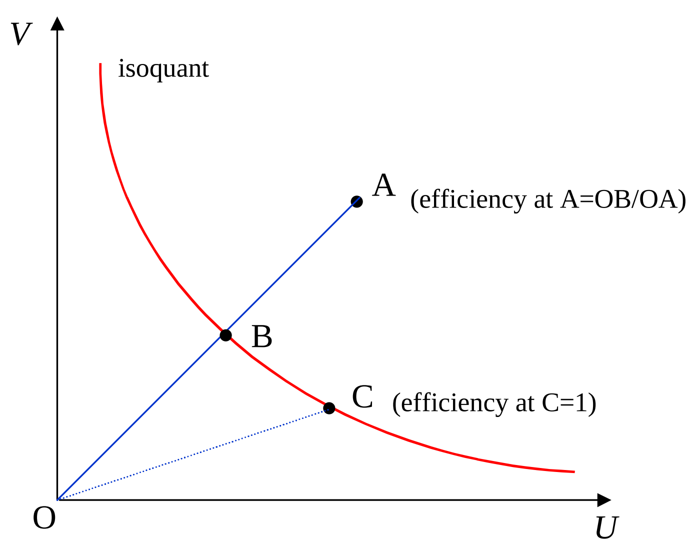

Figure 1: The efficiency frontier in the V-U space

By definition of the efficiency frontier, all observed input combinations will be (at given output level) either on or to the right of an isoquant curve like the one depicted in figure 1 (efficiency cannot be larger than one). In the specific case where function $F(.)$ in equation (1) is Cobb-Douglas, a standard representation of the stochastic production frontier model would then be:

$$\log H_{it} = [\alpha + \beta_1 \log V_{it-1} + \beta_2 \log U_{it-1} + \upsilon_{it}] + \log e_{it}\,, \quad \text{subject to } \; e_{it} \leq 1. \tag{4}$$

The term in between square brackets corresponds to (the log of) function $F(.)$. The position of this frontier is affected by a random term υ_{it},which we will assume to be iid $N(0, \sigma_v^2)$. The efficiency parameter e_{it} is constrained to be smaller than or equal to one. In many stochastic frontier models this is done by specifying $\log e_{it}$ as a stochastic variable with half standard normal distribution, as in Aigner et al.(1977). In our case, it seems more appropriate to allow the expected value of e_{it} to vary through time and space and be a function of observed characteristics. Efficiency can indeed be seen as the product of two factors, one measuring the rate at which job-seekers and employers meet, the other measuring the probability that a contact leads to a successful match (see for instance van Ours (1991), Anderson and Burgess (2000)). The former factor reflects firms/workers search channels and intensities; it is affected by variables like the replacement ratio, vacancy costs, unemployment duration, age and gender, etc... The latter depends on firms and workers choosiness and is affected by the same variables, plus all the variables that determine skill requirements and the degree of correspondence between workers and jobs characteristics, like firms size and type of activity (manufacturing vs services, e.g.), *etc...* As illustrated in appendix 1, such heterogeneity effects can be taken into account, in the log-linear Cobb-Douglas case, by adding to the initial model linear effects of the proportion of each worker or vacancy type in the corresponding total stock.

Let Z_{it} denote the vector of variables meant to capture heterogeneity effects in region i at time t. The initial stochastic production frontier model (4) is then expanded into:

$$\log H_{it} = [\alpha + \beta_1 \log V_{it-1} + \beta_2 \log U_{it-1} + \upsilon_{it}] + [Z_{it}\, \delta + \epsilon_{it}] \tag{5}$$

where the second square bracketed term corresponds to $\log e_{it}$. To impose the constraint that e_{it} be smaller than or equal to one, we follow Battese and Coelli (1995) and define the random term ϵ_{it} by the truncation of the normal distribution with zero mean and variance σ^2, with truncation point at $-Z_{it}\,\delta$, that is we impose:

$$\epsilon_{it} \leq -Z_{it}\,\delta\,. \tag{6}$$

The parameters of the stochastic frontier and of the efficiency term can be estimated jointly by maximising the log-likelihood of the model. One can next obtain conditional estimates of the efficiency coefficient by computing:

$$\hat{e}_{it} = \mathrm{E}\left[e^{Z_{it}\,\hat{\delta}+\hat{\epsilon}_{it}}|H, V, U, Z\right]\ . \tag{7}$$

A most interesting feature of this stochastic frontier approach is the distinction made between the random terms associated respectively to the production frontier and to the efficiency term. The (estimated) residual $\hat{\epsilon}_{it}$ measures the part of the efficiency differences (across regions and time) that cannot be explained by the Z variables. Because the likelihood of the model can be expressed in terms of $\sigma_T^2 \equiv \sigma_v^2 + \sigma^2$ and $\gamma \equiv \sigma^2/\sigma_T^2$, the value of γ conveniently summarises the relative importance of the residual associated to the efficiency term.

Alternative modelling strategies often used to evaluate matching or production efficiencies obtain as special cases of this fairly general stochastic frontier model. The simple half-normal distribution specification (no Z variables) is obtained by setting $\delta \equiv 0$. The model with stochastic individual effects corresponds to the case where Z is a set of regional dummies and no truncation is imposed (e_{it} not constrained to be smaller than or equal to 1). The "fixed effect" model is obtained by furthermore imposing that the residual associated to the efficiency term is strictly equal to zero ($\sigma \equiv \gamma \equiv 0$).

3 Data

We estimate the stochastic frontier model (5) on French panel data, covering twenty-two regions over the period 1990:3 till 1995:2. Details on data sources and definitions are given in appendix 2.

V_{it-1} is defined as the stock of all unfilled vacancies registered with the public employment agency (ANPE) at the end of month $t-1$. It includes part-time and temporary jobs as well as permanent full-time jobs, but excludes special vocational training programmes (the so-called "*stages*"). Composition effects (the share of full-time permanent job offers has dramatically decreased over time) will be taken into account via the Z variables. U_{it-1} measures the stock of unemployed workers looking for a permanent full-time job at the end of month $t-1$ and registered with ANPE. Unemployed workers who enter special training programmes are *not* included. Although most job seekers registered with ANPE look for full-time permanent jobs (more than 93% of them), many of them eventually accept a temporary or part-time job (in February 1995, only one third of all filled vacancies were permanent full-time jobs). Including part-time and temporary jobs in the definition of V_{t-1} seems from this point of view a most natural choice. Finally, H_{it} measures the flow of unemployment-to-job transitions. H_{it} is defined as the number of unemployed job seekers in region i who found a job in period t (again excluding special training programmes)[1].

The across region average values of labour market tensions (measured by V/U) and of the proportion of unemployed workers who find a job during the month (H/U) are reproduced in the left panel of figure 2. Not surprisingly, both variables display wide seasonal fluctuations. From 1990:3 till 1993:2, there is a clear downward trend in the probability of finding a job (as measured by H/U). There is no similar trend for labour market tensions, at least after 1990. The contrast between the two suggests a declining "matching efficiency". The right panel of figure 2 gives some information about regional differences. It illustrates the correlation observed between (the log of) hiring probabilities ($\ln(H/U)$) and (the log of) labour market tensions ($\ln(V/U)$) across regions in February 1995. There is a slightly positive correlation and a slope coefficient approximately equal to 0.37. The regions that are above the line look *a priori* more efficient. This exercise of course remains much too simple. It imposes constant returns to scale; it fails to exploit the time-series information and to control for the many factors which may

[1]We neglect interregional migration, which seems to play a modest role in France, at the region level (see Petrongolo and Wasmer (1999)). We also neglect the impact of on-the-job search. This is compatible with our definition of hirings -which does not include job-to-job flows-; it may however bias the parameter estimates if employed workers compete with unemployed ones for the vacancies posted at the national agency. See for instance Anderson and Burgess (2000), Broersma and van Ours (1999). Interregional comparisons will remain unbiased as long as the relative importance of on-the-job search does not vary across regions or is accounted for by the turnover variable.

affect matching efficiency (the Z variables).

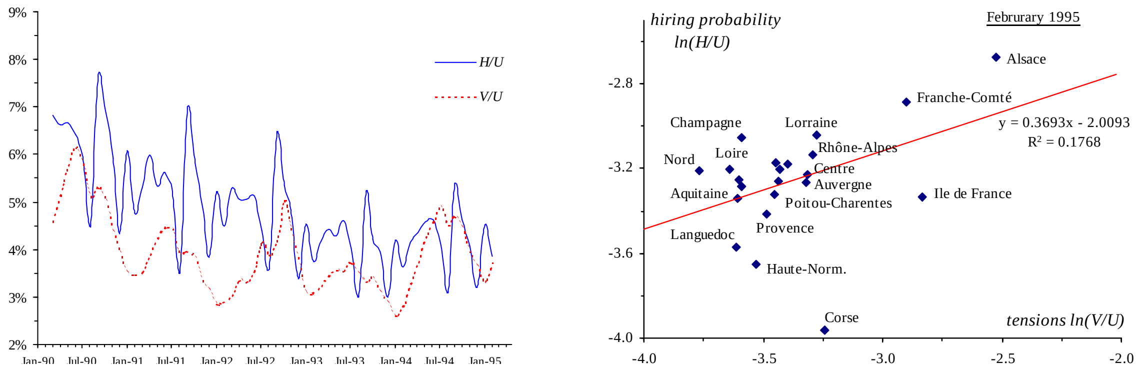

Figure 2: Labour market tensions (V/U) and hiring probabilities (H/U) over time (left; national averages) and across regions (right; February 1995)

To explain matching efficiency differences both over time and across regions, we thus introduce variables meant to capture the characteristics and behaviours of firms and unemployed job seekers. Among unemployed job seekers, we distinguish young workers (< 25), older workers (> 50), immigrants, skilled workers (defined by their former occupation: supervisors, technicians and managerial workers), women and long-term unemployed workers (>1 year). The size of each group is measured in percentage points of total unemployment (see appendix for a brief justification). Differences between these groups may reflect different search intensities, willingness to accept received job offers (willingness to accept a temporary job, e.g.) and/or firms' attitudes (ranking, discrimination). Effects coming from firms size or type of activity will be proxied by the turnover rate and the proportion of permanent (as opposed to temporary) job offers. Firms behaviour may also change with growth (both across regions and over the cycle), for instance because growth may change the relative cost of screening job applicants. Such changes may affect short- and long-term unemployed workers differently (see for instance Lockwood (1991)). The net employment growth variable (lagged one period to avoid tautological effects) is meant to capture this type of influence.

The setup of special training programmes may also affect aggregate matching efficiency even though the unemployed workers entering these programmes are excluded from the definitions of H_t, V_{t-1} and U_{t-1}. To the extent that these special programmes are in effect targeted on workers with lower employment prospects, removing them from the market will automatically increase the observed average matching efficiency (see appendix 1). We control for such effects by introducing among the Z variables the proportion of unemployed workers going into special training programmes. Finally, population density is meant to capture effects coming from the density of economic activities and the probability that a contact is established between the right employer and employee. The same variable may of course capture different effects simultaneously. For instance, the proportion of long term unemployment may capture both business cycle effects and more structural difficulties (firms behaviour, disenfranchisement and/or skill or occupational mismatch e.g.). To the extent that services will be relatively less important in rural areas e.g., the population density variable may capture effects related to the type of economic activity as well as effects coming from "market thickness".

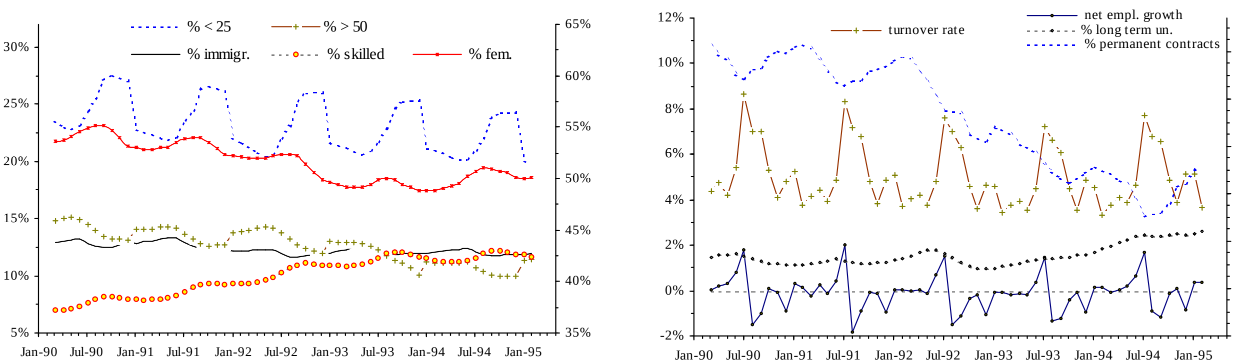

Figure 3: Unemployed worker characteristics and other factors potentially correlated with efficiency; the variables "*% fem. workers*" -left panel-, "*% long term un.*" and "*% permanent contracts*" -right panel- are measured on the right scale; all values represent national averages

The national average values of the most relevant variables are reproduced in figure 3. Some of these variables have pronounced seasonal fluctuations. A few features are worth pointing out. The proportion of both young and older workers in unemployment decreases over time, which implies that the proportion of middle-aged unemployed workers increases. The proportion of unemployed female workers is first decreasing, next increasing. Quite surprisingly, the proportion of skilled workers has been steadily increasing in the early nineties (see Goux and Maurin (1993)) and remained high afterwards. The proportion of immigrants remains stable at the aggregate level (there may of course be substantial differences across regions). Finally, we observe after

1992-93 an increase in the share of long term unemployment and a pronounced decrease in the proportion of unfilled vacancies offering a permanent contract.

4 Results

The model is estimated with the *Frontier* programme of Coelli (1992), under the assumption that the residuals are *iid*. Monthly dummies are added alongside the constant term of the matching function, to capture the effects of purely seasonal fluctuations in the flows of hirings and the stocks of vacancies and job seekers. For coherency, we must eliminate the purely seasonal components that may be present in the Z variables. This was done (for each region separately) by first regressing each Z variable on a time trend and monthly dummies, and next eliminating the monthly effects when significant. The Z variables also include a constant term[2] (on top of the one included in the frontier definition; see equation (5)) and a time trend common to all regions. We further allow for non-linear population density effects by including a quadratic term.

Parameter Estimates

The main estimation results are reported in table 1. *Model 1* gives the results obtained by unrestricted estimation over the entire sample (22 regions from 1990:3 till 1995:2). The elasticities of hirings with respect to unfilled vacancies and unemployment are both positive and quite significantly different from zero. The unemployment elasticity is estimated to be much larger than the vacancy elasticity (0.80 against 0.21). These estimates are in line with those reported in other studies where hirings are defined as unemployment-to-job flows (see table 1 of Broersma and van Ours (1999) *e.g.*). Most variables used to explain efficiency have *t-stat* values well above two, except for the time trend, the proportion of older workers, the proportion of skilled workers and net employment growth (lagged one period), which have small and statistically non-significant effects. The significant variables have a sign that is easily interpretable (a positive sign means a positive effect on efficiency). A larger proportion of young workers, of immigrants and/or of female job seekers increases the probability of matching. This may reflect different degrees

[2] If there were no other Z variable, this constant term would serve to measure average efficiency.

of choosiness, for instance because these groups are more willing to accept temporary jobs, which represent a substantial fraction of total vacancies. In the same vein, one obtains that a larger percentage of permanent contract offers increases matching efficiency. The turnover variable (lagged one period) has a positive significant effect, which may capture sectoral effects (firms size, type of activity and work organisation, ...). Long term unemployment has a negative effect, as expected. Special training programmes have, via their effect on the composition of the stock of job seekers, a positive effect on aggregate efficiency, which suggests that these programmes are (as one would expect) effectively targeted on unemployed workers with below average matching efficiencies . As in Coles and Smith (1996), geographic density seems to affect hiring probabilities. The estimated population density coefficients imply a concave relationship between efficiency and density. Except for the Ile de France region where the estimated effect is close to zero[3], population density has a positive, sizeable effect on efficiency: population density is estimated to generate an efficiency increase of about 33% in the Nord-Pas de Calais region (the most densely populated region after Ile de France) in comparison with the least densely populated areas.

The estimated vacancy and unemployment elasticities of *Model 1* imply barely increasing returns to scale (1.01). The constant returns-to-scale restriction is easily accepted by a likelihood-ratio test and leaves the parameter estimates almost unchanged (see *Model 2*). It is worth pointing out that removing the quadratic term from the specification of the density effect leads to a return-to-scale coefficient estimate that is statistically significantly larger than 1. Increasing returns to scale imply that *ceteris paribus* larger regions should have larger hiring-to-unemployment ratios; if not, they would be characterised by lower matching efficiencies. Because the size (U and V) and density variables are strongly correlated (for instance the Ile de France region is by far the largest region -19% of total (un-)employment- and also the most densely populated one -about seven times the average density-), larger estimates of the return-to-scale parameter may well compensate an inappropriate representation of the density effect. In any case, disentangling return-to-scale and density effects is potentially problematic, but in our case turns out to have

[3]Population density (hundreds of inhabitants per square kilometre) reaches its lowest values in Corse (0.29), Limousin (0.42) and Auvergne (0.51), its largest values in Alsace (1.98), Nord-Pas-de-Calais (3.20) and Ile-de-France (9.01). Its average value is (1.3).

	Model 1	*Model 2*	*Model 3*	*Model 4*	*Model 5*
	Unrestricted 1990:3-1995:2	CRS restr. 1990:3-1995:2	Ile de France excluded	Sub-period 1993:3-1995:2	Preferred 1990:3-1995:2
constant	−1.9368 (0.1210)	−1.8621 (0.0629)	−2.1179 (0.1218)	−2.5280 (0.1564)	−1.8852 (0.0545)
Vacancies (logarithm)	0.2096 (0.0208)	0.2020 (0.0183)	0.2241 (0.0216)	0.3759 (0.0276)	0.1948 (0.0167)
Unemployment (logarithm)	0.7995 (0.0188)	0.7980 –	0.8057 (0.0178)	0.6975 (0.0239)	0.8052 –
+ 11 *monthly dummies*					
constant	−3.0376 (0.2621)	−3.1158 (0.2379)	−3.0537 (0.2535)	1.9891 (0.8835)	−3.0211 (0.1754)
trend	0.0009 (0.0009)	0.0009 (0.0008)	0.001 (0.0009)	0.0198 (0.0042)	
% < 25 years[a]	0.0285 (0.0042)	0.0291 (0.0042)	0.0342 (0.0052)	0.0619 (0.0128)	0.0308 (0.0031)
% > 50 years[a]	−0.0012 (0.0098)	−0.0001 (0.0063)	0.0029 (0.0064)	−0.0279 (0.0208)	
% immigrants	0.0145 (0.0027)	0.0146 (0.0026)	0.0147 (0.0034)	−0.0141 (0.0086)	0.0138 (0.0024)
% skilled workers	−0.0051 (0.0037)	−0.0036 (0.0032)	0.0013 (0.004)	0.0277 (0.0104)	
% female workers[a]	0.0300 (0.0037)	0.0302 (0.0038)	0.0263 (0.004)	−0.0622 (0.0194)	0.028 (0.0031)
% long-term unemployment	−0.0140 (0.0024)	−0.0134 (0.0023)	−0.0128 (0.0026)	−0.0512 (0.0073)	−0.0127 (0.0021)
% permanent contracts	0.0017 (0.0007)	0.0016 (0.0006)	0.0024 (0.0007)	0.0036 (0.0022)	0.0015 (0.0005)
turnover rate[a]	0.0461 (0.0079)	0.0474 (0.0067)	0.0505 (0.0069)	0.0631 (0.0197)	0.0453 (0.0057)
net employment growth[a]	−0.0192 (0.0121)	−0.0191 (0.0125)	−0.0158 (0.0151)	−0.0053 (0.0346)	
% special training programmes[a]	0.0813 (0.0152)	0.0809 (0.0155)	0.0726 (0.0153)	0.0335 (0.0174)	0.0772 (0.0108)
population density	0.1569 (0.0257)	0.1639 (0.0199)	−0.0908 (0.0474)	−0.2144 (0.1005)	0.1564 (0.0190)
squared pop.density	−0.0166 (0.0026)	−0.0172 (0.0019)	0.0522 (0.0110)	0.0191 (0.0095)	−0.0167 (0.0019)
σ_T^2	0.0269 (0.0012)	0.0269 (0.0012)	0.0260 (0.0013)	0.0184 (0.0018)	0.0275 (0.0013)
γ	0.5915 (0.1496)	0.5778 (0.1492)	0.5959 (0.1564)	0.2616 (0.0661)	0.6428 (0.1236)
log-likelihood	545.32	545.03	540.82	341.12	536.99 [b]

[a] Purely seasonal fluctuations are eliminated by regression on monthly dummies

[b] Compared to *Model 2*, *Model 5* includes five zero restrictions on the monthly dummies, and four on the efficiency coefficients

Table 1: Estimation results (dependent variable: log of hirings; standard errors between parentheses)

little effect on the other parameter estimates.

Model 3 in table 1 shows that when the *Ile de France* region is excluded from the sample, the return-to-scale coefficient increases slightly to 1.03. Reestimating the constrained model shows that the constant return-to-scale hypothesis is rejected at the 5% level but accepted at the 1% level. Most parameter estimates remain essentially unchanged though. The population density effect has the same amplitude as before. With the Ile de France region excluded, the effect remains positive over the entire range of values considered, this time with a marked convex rather than concave shape.

We pointed out when presenting the data that some of them have quite different patterns before and after 1993. To check for parameter stability, we reestimated the unrestricted model over the subperiod 1993:3 till 1995:2 (which reduces the number of observations from 1320 to 528). The results are presented as *Model 4* in table 1. There are two main changes: (i) a significantly positive and sizeable trend effect, combined with a lower unemployment elasticity (the number of unemployed job seekers increases steadily over the entire sample period); (ii) a larger vacancy elasticity implying significantly increasing return-to-scale, combined with a significantly negative density effect. These findings seem to reflect a multicollinearity problem that becomes especially acute in our reduced sample, as well as the difficulty to explain the low matching efficiency observed in the Ile de France region, which is likely to bias the returns-to-scale and population density coefficient estimates. Except for the proportion of female workers, the other explanatory variables have coefficient estimates that broadly support the earlier findings. Standard errors are of course substantially larger.

Matching Efficiency Estimates

These parameter estimates can be used to compute the conditional expectation of the efficiency coefficient e_{it}. We use the parameter estimates reported as *Model 5* in table 1, that is, after imposing on *Model 1* the constant return-to-scale and all the zero restrictions that are accepted by the data. The distribution of efficiency scores across regions is illustrated in figure 4.

The across-region weighted average efficiency values obtained every year (twelve month average starting in March of every year) are reproduced in the left panel of figure 5. There is a continuous

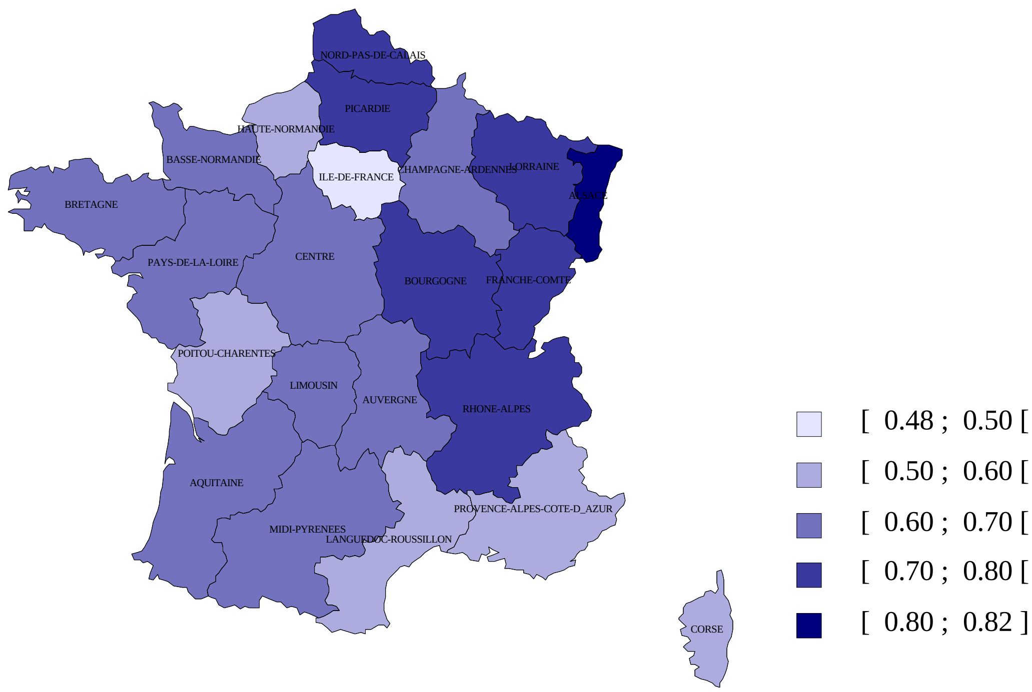

Figure 4: Regional distribution of efficiency scores (average over all observations)

decline in average matching efficiency. From 1990 till 1994, average matching efficiency decreased from 0.74 to 0.54. There are of course large cross-regional differences can be. In 1994, matching efficiency values ranged from a minimum of around 0.42 in Ile de France to a maximum of around 0.75 in Alsace. The size of the one standard deviation interval is indicated on figure 5.

The right panel of figure 5 gives more information on the ranking of regions. It also illustrates the relationship between efficiency and unemployment performance. The figure is drawn for February 1995. Similar figures can be constructed for other periods. The ranking of regions remains actually extremely stable over time (the correlation between the 1990 and the 1995 rankings is 0.74)[4]. The figure suggests, as one would expect, a negative relationship between matching efficiency and unemployment (the absolute value of the correlation coefficient is 0.51, rising to 0.70 when Ile de France and Corse are not included).

[4]The five regions with the best ranking (in terms of gross efficiency averaged over all periods) are (starting with the best): Alsace, Franche-Comté, Nord-Pas de Calais, Lorraine, Rhône-Alpes. The five regions with the worst average ranking are (starting with the worst): Ile de France, Haute-Normandie, Languedoc, Corse, Provence.

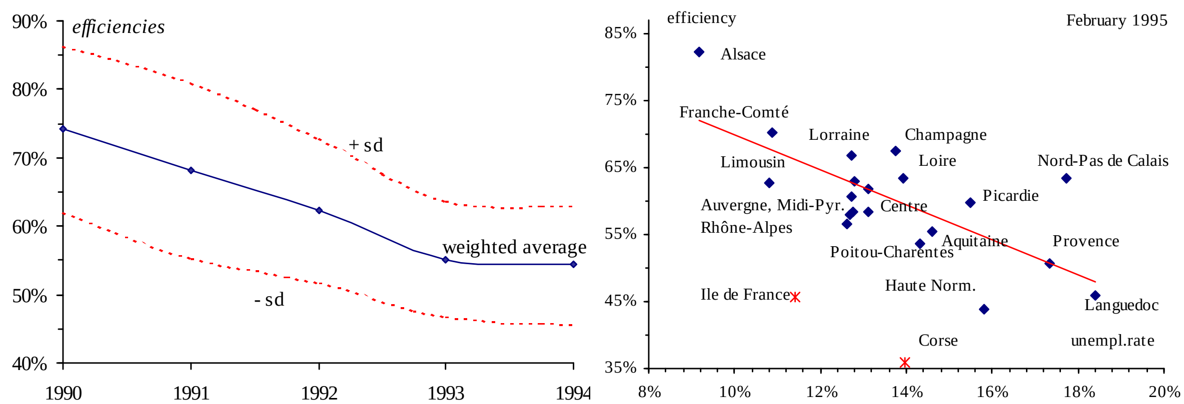

Figure 5: Efficiency estimates across time (left) and regions (right; February 1995)

In order to evaluate how much the Z variables contribute to explain estimated efficiency changes (both over time and across regions), we use in table 2 the concepts of gross *vs* net efficiency. *Gross efficiency* refers to the efficiency estimate obtained when all Z variables take their actual values; *net efficiency* refers to the efficiency estimate obtained when the Z variables are set at fixed, reference values. In our case, these reference values are the average values computed across regions during the first twelve months of the sample (90:3-91:2), that is, the reference values are computed over the best sub-period of the sample in terms of efficiency. The results are reported in table 2. The first two columns reproduce the (weighted) average gross and net estimates respectively. The top part of the table gives the average efficiency scores computed across regions over twelve months subperiods, starting in March 1990 (to simplify the notation, 1990 stands for the period March 1990 till February 1991, and so on for 1991 till 1994). The bottom part of the table gives the averages computed across time for each region separately, starting with the most efficient one (Alsace). One easily checks that the ranking of regions is about the same for both gross and net efficiency (rank correlation coefficient 0.99). This suggests that the regions with the best characteristics (higher $Z\delta$ values) are also those who are "intrinsically" performing better in terms of matching efficiency (have higher mean ϵ values), as one would expect if the most dynamic regions are also those who have a larger proportion of young workers and/or a lower proportion of long term unemployment *e.g.*

Columns (3) to (5) report the difference in percentage points between gross and net efficiency estimates. A negative (resp. positive) value implies that the Z variables, compared to their reference values, have had a negative (resp. positive) impact on the efficiency score of the specific sub-period or region considered. For instance, the difference between the reference and the 1994 values of the Z variables has had a negative impact on matching efficiency equal to -6.4%. This value is a weighted average across regions. The impact of the Z variables may of course differ substantially from region to region. In the worst hit region, the difference between gross and net efficiency was as low as -8.5% in 1994 (column (4)); in the best region, the difference was limited to -2.3% (column (5)). The bottom part of the table gives a similar information for each region separately. For instance, the Z values observed in Picardie, in comparison with the 1990 reference values, implied on average a -5.5% efficiency disadvantage, fluctuating between -12.5% and $+0.6\%$ over the entire sample period.

The efficiency frontier is determined by the situation observed during the best periods (essentially the first twelve months of the sample) in the best regions (mainly Alsace). The part of the efficiency decline (or alternatively the part of the increased distance to the frontier) due to changes in the Z variables can be measured by comparing gross and net efficiency changes. Gross efficiency differences (across time or regions) include both the effects coming from differences in the Z variables *and* those coming from unexplained factors (the residual term ϵ_{it}); net efficiency differences measure only the contribution of unexplained factors. Subtracting the two can thus be interpreted as a measure of the contribution of the Z variables to the estimated efficiency differences, across time or regions. These measures, expressed in percentage points of the gross change, are summarised in column (6) of table 2, for the temporal and regional dimensions separately. We see from the table that changes in the Z variables contribute to explain about 17% of the estimated efficiency changes between 1990 and 1991, and 24% of the decline observed between 1990 and 1994[5]. A similar approach is used for the spatial dimension. The reference is here given by the Alsace region. From the table, we can see that more than half the difference between the matching efficiencies of Alsace and *e.g.* Picardie can be related to changes in the

[5]For instance, the value corresponding to 1994 is obtained by computing over every region separately (i) the difference between the 1990 and 1994 gross efficiency values, (ii) next subtracting from it the difference between the 1990 and the 1994 net values, (iii) expressing the result in percentage points of (i), and finally (iv) computing the weighted average over regions. Applying the same computation on weighted average gross and net efficiency values directly would yield $100 * [(0.74 - 0.54) - (0.76 - 0.61)]/[0.74 - 0.54] = 25.0\%$, instead of the 24.0% reported in the table. The values for each region -bottom part of the table- are computed in a similar fashion, with Alsace as the reference.

	Efficiency Est. (averages)		Differences Gross-Net (in % of gross efficiency)			Explanatory power of Z-variables (in %)	
	Gross (1)	Net (2)	Aver. (3)	Min. (4)	Max. (5)	Frontier distance[a] (6)	$(\sigma_g - \sigma_n)/\sigma_g$ (7)
Years (weighted averages over all regions)							
1990	0.74	0.76	−1.6	−5.8	4.0	–	21.6
1991	0.68	0.71	−2.6	−5.9	3.5	17.0	18.7
1992	0.62	0.67	−4.3	−7.1	1.6	22.1	19.0
1993	0.55	0.61	−5.6	−7.7	−0.7	20.9	17.0
1994	0.54	0.61	−6.4	−8.5	−2.3	24.0	14.3
Regions (averages over all periods)							
Alsace	0.82	0.80	1.3	−5.3	6.6	–	
Franche-Comté	0.78	0.79	−2.6	−10.5	5.5	69.3	
Nord - Pas de Calais	0.76	0.78	−3.5	−10.9	1.0	63.2	
Lorraine	0.74	0.76	−3.8	−10.9	2.6	48.0	
Rhônes-Alpes	0.73	0.75	−2.7	−10.9	5.0	32.2	
Picardie	0.73	0.77	−5.5	−12.5	0.6	54.3	
Bourgogne	0.71	0.74	−3.7	−11.7	3.0	32.9	
Bretagne	0.70	0.72	−4.3	−10.2	0.9	33.1	
Champagne	0.69	0.73	−5.8	−13.0	0.7	39.1	
Pays de la Loire	0.67	0.70	−5.1	−12.6	2.1	30.2	
Basse-Normandie	0.66	0.69	−4.8	−12.1	1.3	26.4	
Centre	0.64	0.67	−5.0	−12.6	1.9	23.5	
Midi-Pyrénées	0.63	0.68	−8.3	−15.6	−0.2	33.9	
Aquitaine	0.63	0.67	−6.9	−12.6	0.1	28.4	
Auvergne	0.62	0.67	−8.7	−13.5	−3.0	32.6	
Limousin	0.60	0.64	−6.8	−10.9	−1.6	23.6	
Poitou - Charentes	0.59	0.64	−9.6	−16.8	−2.1	28.7	
Provence	0.58	0.63	−10.0	−16.7	−2.0	27.8	
Corse	0.56	0.62	−12.3	−18.4	−2.3	30.3	
Languedoc	0.53	0.59	−10.8	−17.5	−1.8	23.5	
Haute-Normandie	0.53	0.56	−7.7	−15.7	1.0	16.9	
Ile-de-France	0.48	0.56	−14.9	−22.7	−7.9	24.8	
All							
	0.63	0.67	−7.8	−22.7	6.6	31.5	

[a] Percentage of efficiency slacks measured with respect to the year (1990) or the region (Alsace) closest to the frontier

Table 2: Estimated Efficiency Coefficients

Z variables. Averaging over time and regions gives an average contribution of the Z variables to estimated efficiency differences equal to 31.5%. The values reported in table 2 are illustrated in figure 6. The last column of table 2 provides an alternative means of evaluating the role of the Z variables in explaining cross-regional differences. The values reported are the difference (in percentage points) between the standard deviations of gross and net efficiencies, computed across regions from 1990 till 1994. This comparison shows that about 15% to 20% of the standard deviation of gross efficiencies across regions can be related to cross-regional differences in the values of the Z variables.

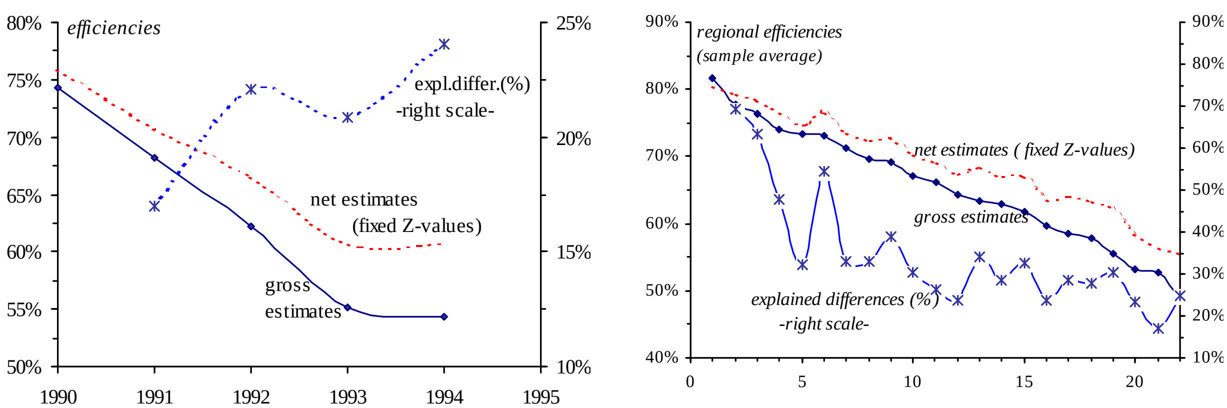

Figure 6: Gross and net efficiencies (left scale) and contribution of the Z variables to estimated efficiency differences (right scale), across time (left panel) and across regions (right panel)

Table 3 gives information about the individual effects of the most important Z variables. The first three columns indicate the values taken by the variables: reference values, range of variation in both the time and the regional dimensions. Not surprisingly, most of the variability is observed in the regional dimension. The next three columns measure the contribution of each Z variable to efficiency. Column (4) gives the marginal effect of each variable, calculated at sample mean (see Frame and Coelli (2001)). The marginal effect measures the efficiency change (in percentage points) obtained by increasing the corresponding Z variable by one percentage point. Columns (5) and (6) give for each Z variable the total efficiency effect of the variation indicated in columns (2) and (3), for the time and the regional dimension respectively. Because the effects of Z changes are non-linear, the values reported in the last two columns are not simply the marginal effect times the corresponding Z variation.

	Explanatory Variables			Efficiency Effects		
	reference	max. variation		marginal	max. effect	
	value (1)	years (2)	regions (3)	effect (4)	years (5)	regions (6)
% youngsters (< 25)	29.2	−2.8	13.5	0.67	−0.9	−7.4
% immigrants	10.1	−1.5	27.3	0.30	−0.4	−6.5
% female workers	55.4	−3.8	15.9	0.61	−1.5	−5.9
population density	1.3	0.0	2.9	3.41	−	−4.5
% long term unemployment	31.2	4.6	14.8	−0.28	−0.8	−3.4
turnover rate (%)	5.0	−0.4	6.0	0.99	−0.3	−3.2
% sp. training programmes	1.5	−0.3	1.3	1.68	−0.2	−1.4
% permanent contracts	69.9	−28.4	30.6	0.03	−0.6	−0.5

Table 3: Contributions of individual Z variables to estimated efficiency

These results are illustrated in figure 7. Because there is a continuous decrease in estimated efficiency from the beginning to the end of the period covered by our analysis, variations along the time dimension correspond to those observed or estimated from 1990 till 1994. In the time dimension, all Z variables (excluding the density variable) have more or less similar effects on efficiency (around one percentage point). The effect is somewhat larger for the proportion of female workers, and quite small for special training programmes despite a large marginal effect. It is worth noticing that the huge increase in the proportion of temporary contracts seems to have had little effects on matching efficiency, even though most workers are looking for a permanent job (see above discussion). As for the regional dimension, four variables emerge and seem to play a dominant role: proportion of youngsters, of immigrants and of female workers, and also population density. The effect of the other variables, although smaller, is far from negligeable, except for temporary contracts.

5 Conclusions

Aggregate matching efficiency can be affected by a variety of factors. Our objective in this paper was to examine what can be learnt by using a *stochastic production frontier* approach. To the extent that the matching process is compared to a production process with unemployment and vacancies as inputs, the stochastic production frontier methodology seems a most natural one to examine the determinants of efficiency. In this setup, aggregate matching efficiency

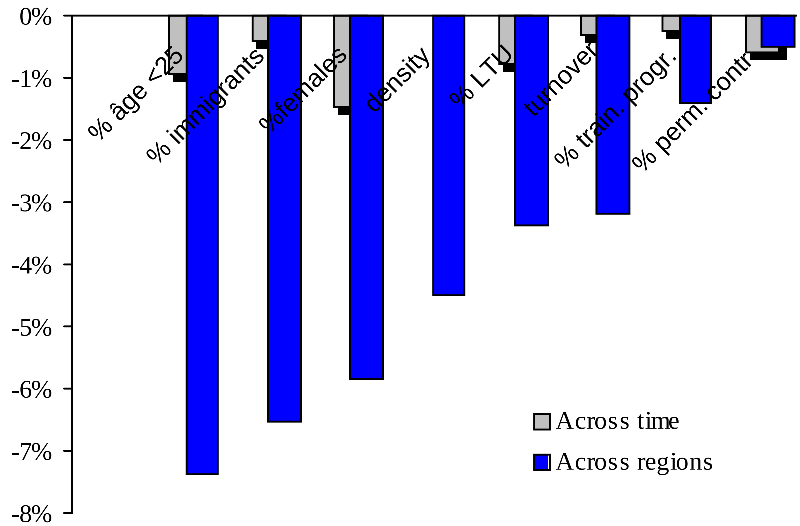

Figure 7: Individual efficiency effects of the Z variables

becomes a stochastic function of variables accounting for the heterogeneity of job seekers and firms. The model is estimated on French panel data covering twenty-two regions from March 1990 till February 1995. Our findings suggest that average matching efficiency has decreased steadily in the early nineties. There are also wide differences across regions. These regional differences in matching efficiency are fairly stable over time and negatively correlated to the regional unemployment rates. On average, about 30% of the variations of efficiency observed across time and regions can be related to changes in the explanatory variables used in the model. The variables that play the most important role are the proportion of youngsters, females and immigrants in the stock of job seekers. Population density seem also to have a significant positive impact on matching efficiency. Population density and scale effects are however difficult to disentangle. Finally, it is worth noticing that the huge decline in the proportion of permanent job offers (from 75% in 1990 to 45% in 1994) has apparently had little effect on matching efficiency, despite the fact that most (registered) job seekers are looking for a permanent job.

Summing up, frontier analysis appears as a promising field for the study of matching efficiency. Frontier analysis not only offers a well-developed measurement methodology but, at the same time, it affords to identify the role of potential explanatory factors inside the "black box".

References

Agullo M.H. (1999) "Efficience du marché du travail: le cas français" in M.Catin, J.Y.Lesueur and Y.Zenou (eds) *Emploi, concurrence et concentration spatiales*, Paris: Economica, 21-38.

Aigner D.J. and C.A.K.Lovell and P.Schmidt (1977) "Formulation and estimation of stochastic frontier production function models", *Journal of Econometrics*, 6, 21-37.

Anderson P.M. and S.M. Burgess (2000) "Empirical Matching Functions: Estimations and Interpretation Using State-Level Data", *The Review of Economics and Statistics*, 82(1), 93-102.

Battese G.E. and T.J.Coelli (1995) "A Model for Technical Inefficiency Effects in A Stochastic Frontier Production Model for Panel Data", *Empirical Economics*, 20, 325-332.

Blanchard O.J. and P.Diamond (1989) "The Beveridge Curve", *Brookings Papers on Economic Activity*, 1-60.

Blanchard O.J. and P.Diamond (1990) "The Aggregate Matching Function", in P.Diamond (ed.) *Growth, Productivity, Unemployment*, Cambridge: MIT Press, 159-201.

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

Coelli T.J. (1992) "A computer program for frontier production function estimation. Frontier Version 2.0", *Economics Letters*, 39, 29-32.

Coles M.G. and E. Smith (1996) "Cross-Section Estimation of the Matching Function: Evidence from England and Wales", *Economica*, 63:252, 589-98.

Debreu G. (1951) "The coefficient of resource utilization", *Econometrica*, 19, 273-292.

EC (1999) *Employment in Europe 1999*, Report of the European Commission, Brussels.

Entorf H. (1998) *Mismatch explanations of European Unemployment. A Critical Evaluation*, European and Transatlantic Studies, Berlin:Springer.

Farrell M.J. (1957) "The measurement of productive efficiency", *Journal of the Royal Statistical Society*, Series A, 120, 253-281.

Frame W.S. and T.J. Coelli (2001), "U.S. Financial Services Consolidation: The Case of Corporate Credit Unions", *The Review of Industrial Organisation*, 18, 229-242.

Goux D. and E.Maurin (1993) "Emploi et chômage des cadres 1990-1992", INSEE Première, No 262, June 1993.

Gorter C. and J.van Ours (1994) “Matching Unemployment and vacancies in regional Labor Markets: An Empirical Analysis for the Netherlands”, *Papers in Regional Science*, 73(2),153-167.

Ibourk A. and S. Perelman (2000) “Frontières d’efficacité et processus d’appariement sur le marché du travail au Maroc”, *Economie et Prévision*, 15/9.

Lockwood B. (1991) “Information externalities in the labour market and the duration of unemployment”, *Review of Economic Studies*, 58, 733-754.

Maillard M. (1997) “L’appariement des emplois vacants et du chômage dans les marchés régionaux du travail en France: Une analyse sur données de panel”, *Revue d’Economie Régionale et Urbaine* 2, 155-170.

Petrongolo B. and C.A. Pissarides (2001) “Looking Into the Black Box: A Survey of the Matching Function”, *Journal of Economic Literature*, 39(2).

Petrongolo B. and E.Wasmer (1999) “Appariement des marchés régionaux du travail: une comparaison France - Grande-Bretagne” in M.Catin, J.Y.Lesueur and Y.Zenou (eds) *Emploi, concurrence et concentration spatiales*, Paris: Economica, 39-53.

Pissarides C.A. (1990) *Equilibrium Unemployment Theory*, Oxford: Basil Blackwell.

Shephard R.W. (1953) *Cost and production Functions*, Princeton: Princeton University Press.

van Ours J.C. (1991), “The Efficiency of the Dutch Labour Market in Matching Unemployment and Vacancies” *De Economist*,139 (3) 358-378.

Warren R.S. Jr (1991) “The Estimation of Frictional Unemployment: A Stochastic Frontier Approach”, *The Review of Economics and Statistics*, LXXIII(2), 373-377.

Appendix 1 Heterogeneity and Aggregate Matching

We consider a non-stochastic model where different groups of job seekers can have different search intensities. The Cobb-Douglas matching function is then written as follows:

$$H_t = \alpha\, V_{t-1}^{\beta_1} \left(\sum_j (1 + c^j)\, U_{t-1}^j \right)^{\beta_2} . \tag{A.1}$$

The c^j coefficients represent deviations from average search intensity, that is, the average value of c^j across job seeker groups is zero; positive (resp. negative) values are associated to groups with above (resp. below) average search intensity. If all groups had identical search intensity, then $c^j = 0\ \forall j$ and we would be back to the standard model without heterogeneity. Rearranging the terms yields:

$$\begin{aligned} H_t &= \alpha\, V_{t-1}^{\beta_1} \left(U_{t-1} + \sum_j c^j\, U_{t-1}^j \right)^{\beta_2} , \\ &= \alpha'\, V_{t-1}^{\beta_1}\, U_{t-1}^{\beta_2} \left(1 + \sum_j c^j\, \frac{U_{t-1}^j}{U_{t-1}} \right)^{\beta_2} . \end{aligned} \tag{A.2}$$

Taking logs and considering that the term in between brackets is close to 1, we obtain:

$$\log H_t \approx \alpha + \beta_1 \log V_{t-1} + \beta_2 \log U_{t-1} + \sum_j \delta_j \frac{U_{t-1}^j}{U_{t-1}} , \tag{A.3}$$

where $\delta_j \equiv \beta_2\, c^j$. In other words, the total number of matches is a function of the total number of job vacancies and job seekers, plus a set of variables representing the share of each group j in total unemployment. A similar development could be made with respect to job vacancies.

The setup of special training programmes may affect aggregate matching efficiency even though the unemployed workers entering these programmes are excluded from the definitions of H_t, V_{t-1} and U_{t-1}. To the extent that special employment programmes are in effect targeted on workers with lower employment prospects, removing them from the market will increase the observed average matching efficiency. Let us illustrate this point by generalising the previous specification. We denote S_{t-1}^j the number of unemployed workers of group j who enter a special training programme and are withdrawn from the official unemployment statistics. Equation (A.1) now becomes:

$$H_t = \alpha\, V_{t-1}^{\beta_1} \left(\sum_j (1 + c^j) \left[U_{t-1}^j - S_{t-1}^j \right] \right)^{\beta_2} , \tag{A.4}$$

where U^j_{t-1} represents the initial number of type j job seekers; H_t does *not* include flows into special training programmes. Provided U_{t-1} is defined as the total number of unemployed workers *excluding* those who enter special training programmes, *i.e.*, $U_{t-1} \equiv \sum_j \left[U^j_{t-1} - S^j_{t-1}\right]$, equation (A.2) is changed into:

$$\begin{aligned} H_t &= \alpha\, V^{\beta_1}_{t-1}\, U^{\beta_2}_{t-1} \left(1 + \sum_j c^j\, \frac{U^j_{t-1} - S^j_{t-1}}{U_{t-1}}\right)^{\beta_2}, \\ &= \alpha\, V^{\beta_1}_{t-1}\, U^{\beta_2}_{t-1} \left(1 + \sum_j c^j\, \frac{U^j_{t-1}}{U_{t-1}} + \varphi\, \frac{S_{t-1}}{U_{t-1}}\right)^{\beta_2}, \end{aligned} \tag{A.5}$$

where φ represent the weighted search intensity of unemployed workers withdrawn from the market and entering special training programmes, more precisely $\varphi \equiv -\sum_j (S^j_{t-1}/S_{t-1})\, c^j$. If these workers have below average search intensities (negative c^j's), φ will be positive, that is, an increase in the size of special training programmes mechanically increases the average observed matching efficiency.

Taking logs and considering again that the term in between brackets is close to 1, we obtain:

$$\log H_t \approx \alpha + \beta_1 \log V_{t-1} + \beta_2 \log U_{t-1} + \sum_j \delta_j\, \frac{U^j_{t-1}}{U_{t-1}} + \phi\, \frac{S_{t-1}}{U_{t-1}}\,. \tag{A.6}$$

Appendix 2 Data Definitions and Sources

The period covered by the analysis goes from March 1990 till February 1995. The data are published by the French National Institute of Statistics (INSEE). Except for population density (hundreds of inhabitants per square kilometre, regional population estimate of January 1, 1992), all data are available for each of the twenty-two regions of France on a monthly basis.

The stock of unemployed job seekers (U_{it-1}) is defined by the number of unemployed workers registered with the National Employment Agency (ANPE) at the end of the previous month and looking for a full-time permanent job (so-called *category 1* jobs). The flow of hirings (H_{it}) is also obtained from the ANPE and is measured by the number of registered unemployed job seekers who found a job during the month, be it the seeked *category 1* job or a part-time and/or temporary job (so-called *category 2* and *category 3* jobs). To be consistent with this definition of hirings, the stock of unfilled vacancies (V_{it-1}) includes all three categories of jobs registered with ANPE at the end of the previous month. The ratio between the stock of unfilled vacancies of categories 1 and 2 and the total stock of unfilled vacancies defines the proportion of permanent job offers.

The monthly variables describing the regional composition of unemployment (by age, sex, skill, nationality and duration) and the size of special employment programmes (number of unemployed workers entering special employment programmes -*stages* for young or long-term unemployed persons- during the month) are collected by ANPE. These figures relate to *category 1* unemployed workers.

The information about hirings and separations (for various reasons : resignation, retirement, dismissal, end of contract, death) used to calculate net job creations and turnover is obtained from an administrative source called *déclaration des mouvements de main d'œuvre* (DMMO). These administrative forms have to be filled by all private and semi-public firms employing at least fifty workers with a normal job contract. This category includes the so-called *contrats de qualification* and *contrats d'adaptation*; it excludes all interim workers as well as the beneficiaries of special professional training programmes (*stagiaires de la formation professionnelle*).