JYVÄSKYLÄ UNIVERSITY
SCHOOL OF BUSINESS AND ECONOMICS

LABOR MARKET MATCHING WITH APPLICATION TO FINNISH DATA

Economics
Pro gradu
18.5.2014

Author: Ville Vento
Supervisor: Jaakko Pehkonen

JYVÄSKYLÄ UNIVERSITY SCHOOL OF BUSINESS AND ECONOMICS

Author Vento, Ville	
Title Labor market matching with application to Finnish data	
Major Economics	**Type of Work** Pro gradu
Date 5.5.2014	**Number of Pages** 69+4

Abstract

This thesis studies labor market matching and takes use of a matching function, which describes the technology how unemployed job seekers and vacant jobs search, meet and form new employment relationships. The goal of the study is twofold: First is to provide a comprehensive critical review of previous literature covering both micro models as well as empirical applications. Major limitations in empirical studies are imperfect measures of variables as well as inability to account for all relevant factors making results methodology-sensitive. The second contribution is to estimate a matching function for Finland using annual data from 2006-2012 and thus giving the most recent knowledge on the matching process in the regional labor markets in Finland. The study uses both conventional panel data concepts and stochastic frontier analysis allowing differentiation of matching technology from inefficiency. The results are mainly in line with previous studies and indicate that one percent increase in the number of unemployed job seekers increases the outflow into employment by 0,9 percent whereas same increase in the number of vacancies reported at local labor offices only raises the outflow by 0-0,1 percent. The data shows that most vacancies are filled during a year, but they do not seem to help much the local unemployed, as many of those vacancies are filled by non-unemployed job seekers as well as job seekers from other regions. Efficiency in matching has decreased during the research period and it is negatively correlated with the share of long-term unemployment and positively correlated with the shares of unemployed job seekers younger than 25 and older than 55 years old and the share of unemployed in active labor market policies.

Keywords
labour market, matching function, stochastic frontier analysis

Storing location
Jyväskylä University School of Business and Economics

FIGURES

TABLES

CONTENTS

1 INTRODUCTION

Labor markets are in a state of a continuous change. Less productive jobs are being destroyed as workers are reallocated to more productive activities. However, these transitions often do not happen without frictions, that is, factors that prevent instantaneous match of job seekers and vacant jobs. Because of frictions transitions from job to job are often made through a spell of unemployment and time is needed before a new job is found. If there were no frictions, unemployed workers and vacant jobs would match instantaneously.

There are several sources of frictions most of which are very intuitive. Petrongolo and Pissarides (2001, p.390) suggest that frictions stem from information imperfections, heterogeneities, absence of perfect insurance markets, slow mobility and congestion. Since workers and jobs are far from homogenous pools, it takes time to search for a suitable match and negotiate for the terms of employment. This may include waiting for suitable vacancies to open as well as application period and interviews. Moreover, unemployed workers and unfilled vacancies are often spatially dispersed and moving costs, be that of monetary, social or some other form, deter the formation of an employment relationship.

In the basic neo-classical labor market there are no frictions and therefore unemployed workers and unfilled vacancies do not coexist. Naturally, such an economy is an oversimplification of the real world and models based on that assumption are likely to result in inaccurate conclusions. As Yashiv (2007, p.1860) argues, this basic textbook model has problems to account for unemployment as an equilibrium phenomenon, explaining large worker flows as well as explaining some business cycle facts like low cyclicality of real wages. Therefore richer concepts are needed to take into account labor market frictions.

Matching literature explains unemployment as an equilibrium phenomenon stemming from various frictions in actual labor markets. An essential part of the matching literature is a matching function, which has been extensively studied both theoretically and empirically. This concept was originally developed by Diamond and Maskin (1979), Hall (1979) and Mortensen (1982) and further developed by Diamond (1982) and Pissarides (1984, 1985) (Burgess 1993, p.1190).

Matching function gives hires, or matches, as a function of unemployed workers and number of vacancies. It allows us to model the complex reality where workers and jobs differ in many dimensions by a well-behaved and simple function. The matching function makes it possible to add frictions to conventional models with a minimum added complexity. Explicit modeling of frictions would impose enormous complexity into macroeconomic models, and therefore it is practical to treat the matching process as a “black box”, where no reference to the source of frictions is made. (Petrongolo & Pissarides 2001 p. 390.)

The aggregate matching function is a tool in a macroeconomist´s toolbox to be incorporated in macroeconomic models[1] to improve their performance, but it can also be used to examine labor markets per se, which is also the focus of this study. By studying matching one can learn a great deal about unemployment. It has been suggested that unemployment fluctuations are more driven by fluctuations in the transition rate from unemployment to employment than by fluctuations in the separation rate suggesting that the matching process is a key determinant of aggregate unemployment[2] (see e.g. Hall 2005). Empirical studies enable answering such questions as whether demand or supply factors are more important in tackling unemployment or how structural factors contribute to unemployment. It also unveils the extent of externalities that firms and job seekers cause each other and thereby helps us to analyze the efficiency of the search equilibrium. Theoretical micro studies, in turn, help us to design better policies to achieve lower unemployment and inequality. (Petrongolo and Pissarides 2001, p. 391-392, 425.)

Although there is a large body of literature studying the microfoundations of the matching function, there is no consensus on whether the matching function should be of particular form. As Petrongolo and Pissarides (2001) point out, the microeconomic literature has had much more success in suggesting what additional variables should be included in the aggregate matching function, that is, shift variables that do not affect the shape of the function, but determine the location of the curve. This includes variables related to the structure of the stocks like the share of long-term unemployment and factors contributing to the search effort of unemployed workers and firms.

The matching function can be seen as a production function with unemployed workers and open jobs as inputs and hires from unemployment as output. In the empirical applications a Cobb-Douglas specification is often used. Much research has been devoted to study the returns-to-scale of the matching function, which is an interesting question for many reasons. First, the elasticities with respect to input variables can reveal us the importance of demand and supply factors in the matching process. Second, it allows us to assess the efficiency of the matching process related to the size of the labor market. Decreasing returns would imply congestion by firms and workers and thus inefficiency caused by lack of coordination, whereas increasing returns imply support for policies promoting labor mobility. In case of increasing returns-to-scale there may also be several equilibria, some of which contain low and some high activity by firms and job seekers. This indicates that there may be a chance for policy intervention to achieve a more desirable equilibrium. Moreover, Burdett et al. (1993) point out that several theoretical predictions depend on whether there are increasing returns-to-scale or not. Finally, as

---

[1] Yashiv (2007, p.1860) discusses the merits of matching and search models in macroeconomic literature. He uses real business cycle (RBC) model as an example. The neo-classical model is part of the RBC model, but it leads to implausible values of labor supply elasticity needed to produce the observed facts. However, the matching model is compatible with models based on DSGE framework allowing a better performance of those models.

[2] This view has been criticized for example by Davis (2005)

shown by Hosios (1990), a constant returns-to-scale matching function is required to achieve socially efficient labor market equilibrium.

Interestingly, majority of the empirical studies, both at aggregated and disaggregated level, indicate that there exists a matching function that is well approximated by Cobb-Douglas function with roughly constant returns-to-scale (Petrongolo and Pissarides 2001). Elasticity of matches with respect to unemployment is often found to be larger than the one for vacancies suggesting supply side policies to promote employment. However, many matching function studies suffer from methodological problems and therefore caution is needed when interpreting the results.

During the past ten years a growing interest has emerged to study the efficiency of the matching process. As Fahr and Sunde (2002, p.3) argue, it is vital to know what magnitude the inefficiencies are in regional or occupational labor markets in order to design good policies. If elasticities on the stocks are high in certain markets, these inputs have high productivity. However, this high productivity potential may never realize, if there are at the same time high inefficiencies. In such a situation, policies designed towards reducing the inefficiencies are likely to be effective in reducing unemployment. For example, Hynninen et al. (2009) find that eliminating net inefficiencies, that is inefficiency not explained by structural factors, would have lowered the aggregate unemployment rate in Finland by 2.4 percentage points during the sample period.

Recently studies using stochastic production frontier methods in estimating the efficiency of the matching process have been published (See e.g. Ibourk et al. 2004). The novelty of the idea is that inefficiencies are measured as deviations from the potential output, i.e. the maximum output achievable from given inputs, rather than averages. Stochastic production frontier models also have the advantage over traditional panel data methods in that they allow for a more reliable detailed analysis of the sources of inefficiencies. This is also the method used in the empirical study of this thesis.

This thesis surveys the theory and empirical findings of the labor market matching function. It is a very fruitful topic to study, as it allows for drawing important policy implications and gives a deeper understanding on how labor markets actually work. Also the literature is extensive and spans for several decades. Yet the matching approach clearly brings additional insight into the pursuit for understanding labor markets, it also fails at some issues. For example, empirical studies seem to produce systematically different results depending on the methodology used.

The focus of the survey is in micro foundations of the matching process and in a critical review of the empirical applications. The purpose is to create a good understanding on the attempts to model the underlying mechanisms in matching and to discuss weaknesses in the existing empirical literature. For the sake of focus wage setting questions are mainly left aside. The survey is followed by an empirical study using stochastic frontier model in estimation of a matching function for Finland. The aim is to study the role of unemployed job seekers and vacancies in explaining the unemployment-job transitions. The model also reveals us how structural labor force characteristics are connected to efficiency of the matching process. The data set covers years 2006-2012 and thus the results give us the most recent knowledge about matching in the Finnish labor markets. The results confirm findings in the previous literature addressing the importance of supply side in matching.

The structure of this thesis is as follows: The second chapter begins by discussing the Beveridge curve and laying out the basic matching function. After that microfoundations of the matching function and empirical studies are discussed. Chapter four describes the data and the model used in estimating a stochastic production frontier model for Finnish regions and discusses the results. Finally, chapter five concludes.

2 MATCHING FUNCTION

The aggregate matching function is a function that gives hires, or matches, as a function of unemployed workers and available vacancies. It is analogous to a production function in the sense that it "produces" hires as output from unemployed workers and vacancies as inputs. As there are various sources of frictions, it is impossible to add all those aspects into macroeconomic models in a meaningful and technically feasible way. Hence, summarizing the complicated process of heterogeneous agents searching, meeting and negotiating for the terms of employment in a well-behaved aggregate level function makes it easy incorporate frictions into macroeconomic models with little added complexity (Petrongolo and Pissarides 2001 p.390). The model is compatible with most modern macroeconomic models as it is based on optimizing agents, rational expectations and equilibrium outcomes (Yashiv 2007, p.1861)

Blanchard and Diamond (1989, p.3) say that one may legitimately ask whether this kind of simplistic function has any relevance to the reality. Petrongolo and Pissarides (2001, p.391) compare the aggregate matching function to other aggregate functions like production function or demand for money function arguing that the usefulness of those functions ultimately depend on their empirical performance. A large amount of empirical studies conducted by different methods and data sets give surprisingly unanimous view in the favor of the aggregate matching function. Petrongolo and Pissarides (2001, p.393) list four different types of studies from where the evidence for the aggregate matching function has been accrued: Studies considering the joint movement of vacancy rate and unemployment, also known as the Beveridge or UV-curve, estimations of the aggregate matching function with both national and regional data and studying the transition probabilities of individuals from unemployment into employment. The last concept is referred in the literature as hazard rate studies.

As said, the literature on the matching function is very extensive. Although the concept is well applicable to other fields of study as well, the most research is probably done in labor economics. Petrongolo and Pissarides (2001, p.391) suggest that this is due to the great importance that frictions play in labor markets and the existence of data sets that enable empirical estimation of the matching function. Hence, matching function not only serves as tool in macroeconomists´ toolkit to model frictions as a part of conventional macroeconomic models like DSGE models, but can also be used to study the properties of labor markets per se. According to Burgess and Profit (2001, p. 313) the matching approach "has proved to be an essential element in understanding the dynamic processes of labor markets for both labor and macroeconomics".

2.1 The Beverdige curve

The aggregate matching function is firmly connected to the Beveridge curve (see figure 1 below). Petrongolo and Pissarides (2001, p.393) argue that a matching technology given by the standard matching function (1) produces a downward sloping relationship in unemployment-vacancy-space. Beveridge curve, or UV-curve, is an empirical relationship between vacancy rate[3] and unemployment rate. It is usually found to be negative and convex to the origin giving support for the aggregate matching function[4]. Just like the Phillips curve, the Beveridge curve is a steady-state relationship, that is, the curve depicts the combinations of vacancies and unemployment where the inflows into the unemployment are equal to the outflows from it in a given labor market (Ibourk et al. 2004, p.2).

The location of the Beveridge curve is defined by job creation and destruction rates as well as the effectiveness of the matching process (Blanchard and Diamond, 1989 p.2). However, as Bleakley and Fuhrer (1997, p.5) argue, a plot in the UV-space does not directly indicate the efficiency of the matching process. It just depicts the inputs used in the matching process at a given point in time. However, the efficiency of matching process affects the outflow of unemployment and vacancies altering the levels of both over time.

Blanchard and Diamond (1989, p.2) argue that the Beveridge curve is a very useful concept to evaluate labor market dynamics. It can reveal much about the effectiveness of the matching process as well as nature of shocks affecting an economy and thereby be informative about the underlying causes of unemployment. As pointed out by Bleakey and Fuhrer (1997, p.3), the position on the curve can display where the economy is in the business cycle. On the other hand, an outward shift of the Beveridge curve indicates that a rise in unemployment is due to reasons other than lack of demand for labor (Albaek & Hansen 2004, p.516). To summarize, moves along the UV-locus reflect changes over a business cycle, whereas the position of the curve signals structural factors in the given economy. Here a short description of the mechanisms is provided, as the concept is important in understanding the idea behind the matching function. A more detailed analysis on the labor market flows and the Beveridge curve is provided in Blanchard and Diamond (1989).

If the economy is hit by an aggregate demand shock, two simultaneous effects occur. Job destruction and creation rates change in opposite directions causing a change along the UV-locus. For example, in case of negative demand shock more workers are laid off and there are fewer vacancies available. It follows that unemployment increases while vacancy rate falls. In terms of the aggregate matching function, an aggregate demand shock changes the inputs of the matching function but does not alter the technology how matches are formed.

---

[3] Vacancy rate = vacancies / labor force

[4] Lahtonen (2006, p.11) explains that early matching literature often estimated the UV-relationships, as data on flows were rarely available. Since that is has been more common to directly estimate the matching function.

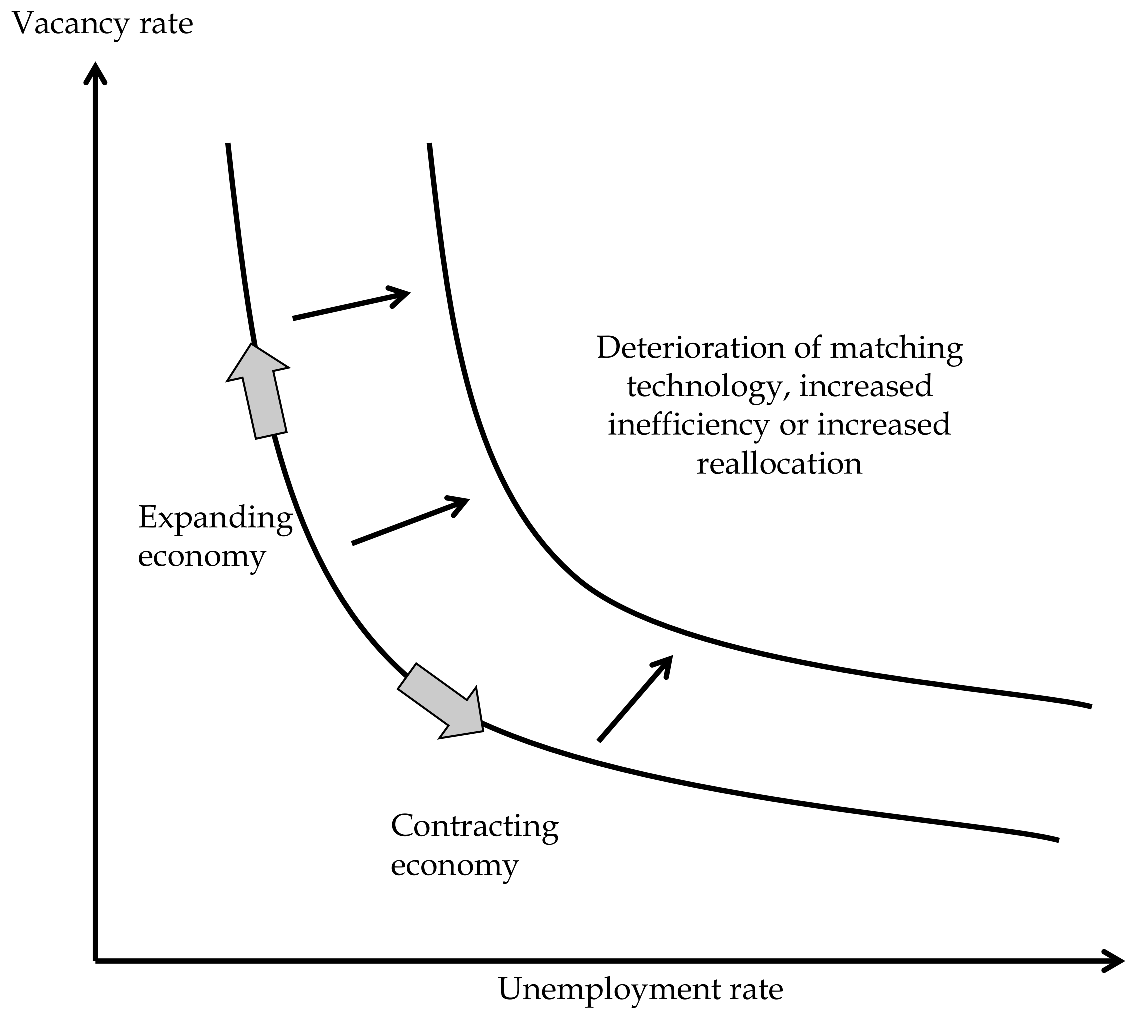


FIGURE 1 The Beverdige curve

Another type of shock that an economy can experience is a reallocation shock. This shock has very different effect from that of an aggregate demand shock. When labor is reallocated from less profitable sectors to more profitable sectors, both job creation and destruction increase shifting the Beveridge curve to the upright. In an ideal economy, effects of a reallocation shock should vanish in the long run as workers shift from declining sectors to growing sectors. However, labor is not often transferable from one sector to another and therefore reallocation shifts may have long-run consequences. This mismatch, be it occupational, regional or skill, may be an important factor explaining the observed shifts of the Beveridge curve. The mismatch hypothesis is addressed more thoroughly in section 1.3.3.

The shifts along the curve or shifts of the curve also bear a close relationship to the question of returns-to-scale in the matching function. If there are increasing returns-to-scale in matching, then increased reallocation should make the matching process more efficient and the impact on unemployment is limited. On the contrary, decreasing returns-to-scale implies more congestion and therefore increased reallocation expectedly increases unemployment more than in the previous case. Also the elasticities on unemployment and vacancies determine the magnitude that aggregate demand shocks affect unemployment. It is easy to see that if the flow from unemployment to employment is more determined by the number of vacancies than number of job seekers, i.e. the elasticity on vacancies is higher, then responses to aggregate demand shocks are larger. On the other hand, if the elasticity on unemployment is higher than the elasticity on vacancies then aggregate demand shocks have a milder impact on unemployment fluctuations. Graphically, the elasticity parameters affect the curvature of the Beveridge curve.

The question whether observed fluctuations in aggregate unemployment are due to aggregate demand shocks or increased reallocation has also been intensively studied. For instance, Blanchard and Diamond (1989, p.50) study post-war data of United States and find that in the short and medium term the fluctuations in unemployment are mainly caused by aggregate demand shocks.

Studies examining the Beveridge curve over time in developed economies often find an outward shift of the curve indicating either increased mismatch or higher degree of reallocation (see e.g. Albaek & Hansen 2004). Bleakley and Fuhrer (1997, p.6) also suggest changes in the labor force size as a potential factor explaining the observed shifts. In the short run, labor force growth leads to an increase in unemployment, as there are more job seekers competing for jobs. In the long run, however, the number of jobs is likely to increase approximately in line with unemployment. Using US data Bleakley and Fuhrer find that combining observed changes in matching efficiency, labor force growth and labor market churning roughly produce the observed shift in the Beveridge curve.

2.2 Basic model

Following Petrongolo and Pissarides (2001) the general form of the matching function can be written as

$$M = m(U,V) \tag{1}$$

where M is the number of hires during a time period, U is the number unemployed workers and V denotes the number of unfilled vacancies. The general assumptions are that the function is increasing in both arguments and concave. That means that matches increase when either the number of unemployed workers or vacant jobs increase, but at a decreasing rate. In equilibrium unemployment theory it is usual to assume that there are constant returns to scale and this hypothesis is also supported by rigorous empirical testing, although there are different results as well. Further assumptions usually made are that m(0,V) = m(U,0) = 0 and if M is the flow into employment during a period and U and V are the stocks at the beginning of a period, then m(U,V) ≤ min(U,V). If constant returns-to-scale to the matching function is assumed, then it is usual to standardize M, U and V by the size of the labor force.

It follows that the average probability that an unemployed worker becomes employed during a given period is m(U,V)/U and similarly the average probability that a vacancy is filled is given by m(U,V)/V. Also, if we can assume stationarity, the mean durations of unemployment and vacancies can be obtained as inverses of those matching probabilities. The heterogeneities can be introduced by making the matching probabilities depend on individual worker or firm characteristics, which is the method used by hazard rate studies. (Petrongolo & Pissarides 2001 p.392.)

The question of the returns-to-scale is an important one. Hosios (1990) derives conditions for social efficient labor market equilibrium and shows that a necessary but not sufficient condition for social efficiency is that the matching technology is homogenous of degree one meaning constant returns to scale. The reason is that to be socially efficient, an equilibrium much be such that the positive and negative externalities cancel out[5]. Constant returns-to-scale also imply that regional agglomeration does not have impact on matching efficiency.

If there are actually increasing returns-to-scale, then there may be many equilibria in the labor market some of which support low and some high activity by firms and job seekers[6] (Kangasharju et al. 2005, p.115). As Petrongolo and Pissarides (2001, p.393) argue, this results from endogenous search effort by both firms and workers. Externalities affect the search decisions of the other input group through a system of feedback effects. In one equilibrium positive externalities make the other side to put more effort into search, which in turn increases the expected gains of search of the other side. Warren (1996) argues that in case of multiple equilibiria there is a chance for welfare-improving policy intervention. Moreover, Burdett et al. (1993) explain the interest in returns-to-scale in the literature by pointing out that many theoretical predictions depend on whether there are increasing returns or not.

On the other hand, if the returns are decreasing, it is a sign that the congestion that firms and job seekers cause to each other dominates the positive effect of more opportunities. Therefore measures to improve the coordination of the matching process may be needed to achieve higher employment.

More formally, following Petrongolo and Pissarides (2001, p.392) if we denote elasticities with respect to unemployment and vacancies η_U and η_V respectively, then η_U-1 exhibits the negative externality, i.e. congestion, of an additional unemployed to other unemployed workers. Similarly, η_V-1 is a measure for negative externality of new vacancies on the existing ones. This phenomenon is intuitive. Increasing other input while keeping the other constant increases the competition within that input, as there are now fewer partners for each seeker to form a match with than before the increase. The externality effect is, however, positive on the other side η_U denoting a positive externality on firms and η_V on job seekers. It follows that the higher the

[5] In addition to constant returns-to-scale it is required that job seekers share of the match surplus equals the elasticity with respect to job seekers.

[6] This is formally shown by Mortensen (1989). He shows that increasing returns-to-scale is a sufficient condition for multiple equilibria if workers and firms share the matching surplus according to efficiency wage or insider-outsider model. (Bunders, 2003 p. 10)

elasticities on input stocks are, the more efficient the matching process is as higher elasticities are related to less congestion and more positive externalities.

2.3 Microfoundations

So far we have treated the matching process as a black box. This means that we have not explicitly modeled the individual level process where unemployed job seekers search for suitable jobs and firms advertise open positions to find workers. Nor have we taken any stance on how matches are formed once firms and job seekers are met. Moreover, we have treated firms and unemployed workers as a homogenous pool, which is naturally an incorrect assumption. In order to understand how matching actually happens at an individual level, what consequences it has to aggregate outcomes and for empirical studies, we must look at theoretical literature. Stevens (2007, p. 848) argues that the reason to study the micro foundations for the matching functions is twofold. First, there is a need to gain a deeper understanding on the nature of frictions and justify the use of the aggregate matching function as a modeling tool. Second, micro theory is needed to determine of what form should the matching function be. While micro studies may be helpful in justifying the use of matching function, they unfortunately fail at giving unambiguous guidance on the functional form.

Again, there is a bulk of literature studying the individual level matching process. A good and more technical literature review is provided by Rogerson et al. (2005). Here the main theories are discussed with emphasis on intuition.

2.3.1 Urn-Ball Model

A good starting point to review the search theory is to view the search process as a process where job seekers and firms meet randomly unaware of each others´ actions. Within this framework the frictions emerge as a result of coordination failure (Hynninen 2007 p.9). As agents with limited resources are not aware of each other´s actions in the absence of coordination, it is possible that they end up crowding some positions while some vacancies receive no applications. This is clearly suboptimal from both individual´s and society´s point of view. The model presented below is known as urn-ball model and it was first studied by probability theorists (see e.g. Butters 1977). Its key features are the incorporation of matching externalities and time-consuming nature of search.

Following Lahtonen (2006, p.20) and Petrongolo and Pissarides (2001, p.401) suppose there are U unemployed workers and V vacancies at the beginning of a period. Moreover, all vacancies and unemployed job seekers are assumed to be identical. In this framework each worker sends only one application and if a firm receives one, it randomly assigns the job to one of the applicants. If a firm does not receive any applications, the vacancy remains open for the next period and similarly unlucky job seekers continue the search in the following period. Hence, even if there are even number of homogenous job seekers and vacancies, it is likely that both coexist.

The probability that a vacancy receives an application is 1/V and similarly the probability that a vacancy remains unfilled is given by $(1-1/V)^U$. It follows that the number of matches during a period is given by a matching function

$$M = V\left[1 - \left(1 - \frac{1}{V}\right)^{U}\right] \tag{2}$$

For a large V and U, the binomial distribution of applications at a given vacancy can be well approximated by a Poisson distribution (Blanchard & Diamond 1994, p.418). Thus, we obtain

$$M = V(1 - e^{-\frac{U}{V}}) \tag{3}$$

It can be seen that this function satisfies the properties of (1) and additionally exhibits constant returns to scale. However, Pissarides and Petrongolo (2001, p. 402) argue that it does not hold in empirical investigations as it implies unrealistic levels and durations of unemployment. In reality, the unemployment durations are clearly higher than predicted by (3).

This framework can be extended to include a few additional friction elements. This way the matching function presented in (3) can be modified to fit the data better. Following Petrongolo and Pissarides (2001, p.402) three different modifications are discussed. Mismatch and different search intensity are discussed in their own chapters, but it is logical to present the models here, as they are extensions to the urn-ball model.

First, consider that workers are unaware which firms have vacancies and randomly choose one to apply. Given N is the level of employment and L stands for the size of labor force, the probability that a vacancy receives no application can be written as $(1 - 1/(N + V))^U$. It follows that the matching function takes the form

$$M = V(1 - e^{-\frac{U}{L-U+V}}) \tag{4}$$

exhibiting increasing returns to scale in U and V. This model also captures some realistic features, as many employers receive open applications that are not related to any specific vacancy. Here again frictions come from imperfect information, but it does not only relate to the actions of other job seekers but also to whether there are job opportunities in a given firm. This element may have relevance even when job seekers are allowed to apply for many jobs simultaneously. Even in that case sending out application for firms where no vacancies exist consumes time and resources available for applying to other jobs.

The second extension allows firms and workers to differ in demand and supply of skills. Denoting the fraction of workers that are suitable for a randomly chosen vacancy as K gives

$$M = V(1 - e^{-\frac{KU}{V}}) \tag{5}$$

Hence, the more heterogeneity there is in skills and firm requirements, the less efficient the matching process is. In this framework where employees lack information about their suitability for jobs, a higher degree of specialization leads to increased mismatch.

We can also allow the search intensity to affect matching outcomes. Suppose that only a fraction of U, denoted by s, applies in each period. This leads to a matching function of the form

$$M = V(1 - e^{-sU/V}) \tag{6}$$

A higher search activity has the same effect as higher K in (6), namely higher efficiency of the matching process, and thus a lower level of unemployment. An alternative interpretation for s, presented by (Stevens 2007), is that s represents the individual search activity rather than the proportion of active job seekers. Furthermore, Blanchard and Diamond (1994, p.418) suggest that s can also be understood to reflect the skill and spatial distributions of workers and jobs. Therefore their model combines the ideas of (5) and (6) sU being the mean of acceptable applications during a period. They also propose that in a more realistic model s should be made dependent on the state of labor market as well as let to vary across workers.

Previously we discussed the search process in discrete time. The concept can also be applied for a continuous time case. Following Stevens (2007, p.848) suppose that job seekers send out applications at a constant rate α. Then model (3) can be written as

$$M = V(1 - e^{-\frac{\alpha U dt}{V}}) \tag{7}$$

Letting dt tend to zero gives a matching rate

$$M = \alpha U \tag{8}$$

This is a linear function of unemployment stock and hence assumes firms to be passively receiving applications with no search activity. A more complex continuous time urn-ball model can also be found in Blanchard and Diamond (1994, p. 419), where vacancies are posted for an exogenous discrete length of time resulting in better technical properties of the model (Stevens 2007, p. 849). A model with search by both unemployed and firms can be found in Mortensen and Pissarides (1999).

To summarize, the urn-ball model is a simple framework to study the effects of lack of information and coordination in the labor market. However, it is not very useful in advising in practical applications and Petrongolo and Pissarides (2001) note that the model has not had much empirical success. Despite that, most empirical matching function studies assume random search, as they do not differentiate unemployed job seekers and vacancies by the duration of the search (Gregg & Petrongolo 2005 p.1988).

2.3.2 A Telephone Line Model

Based on "telephone line" Poisson queuing process originally proposed by Cox and Miller (1965), Stevens (2007) develops a model that under certain assumptions gives microeconomic justification for Cobb-Douglas matching function. The model is also useful in interpreting results of matching function estimations using Cobb-Douglas form. Like the urn-ball model, the model captures both negative and positive externalities from additional vacancies and job seekers. As opposed to urn-ball model, the telephone line model is easy to integrate into standard search models. The process is similar to continuous time urn-ball model, with exception that search is symmetric, i.e. also firms search. The search by firms can be understood as recruitment effort including reading applications and testing potential employees.

The model results in CES-type specification, which is approximately Cobb-Douglas under the assumption of constant marginal search costs, in other words, close to linear cost functions. An important contribution of the model is that it identifies the determinants of unemployment elasticity and elasticity of substitution between unemployment and vacancies. This helps us to interpret empirical values for those parameters.

The idea of the telephone line model is as follows: Let us assume that job seekers send out applications to firms with vacancies at a Poisson rate α. This is analogous to "making calls" to firms. Firms respond to these applications at a Poisson rate γ. These parameters reflect the search and recruitment efforts of workers and firms. When a firm processes an application (answers a call), the value of the potential match is discovered. If the value is high enough, the match is formed. Otherwise, both parties continue search. Furthermore, if a firm receives an application while it is processing another one, the new application will fail. As job seekers are unaware of other job seekers´ actions, similar coordination failure may occur as in the urn-ball model discussed earlier. The congestion effects are easy to see. The more job seekers there are the more likely it is that a firm with a vacancy is already processing an application.

Deriving the matching function is straightforward. Suppose there are V vacancies total and V_0 such vacancies that there is no application being processed. Let U denote the number of job seekers. The number of applications being sent out per unit time is αU and the arrival rate of applications at each vacancy is the ratio of the arrival rate and the number of vacancies $\alpha U/V$. We get the expected number of applications arriving at firms where no application is being processed by multiplying the arrival rate by V_0, that is $\alpha UV_0/V$. The outflow from application processing is the response rate of firms times the number of applications being processed, which is $V-V_0$, hence $\gamma(V-V_0)$. The equilibrium condition can be stated as

$$\alpha U \frac{V_0}{V} = \gamma(V - V_0) \tag{9}$$

Here we can solve the probability that an application will land at a vacancy where no application is being processed, namely V_0/V. We get

$$P = \frac{V_0}{V} = \frac{\gamma V}{\alpha U + \gamma V} \tag{10}$$

Multiplying the probability, or ratio of successful applications, by the number of applications αU yields the number of applications processed at firms.

$$\frac{aU\gamma V}{\alpha U + \gamma V} \tag{11}$$

To derive the matching function, we need to add the probability that a match is acceptable. Let

$$p(z) \equiv \Pr\,[y > z] \tag{12}$$

denote the probability that a processed application leads to a match, y being the productivity and z reservation productivity which equals the sum of the value of being unemployed and the value of having a vacancy. Then multiplying this probability by the number of applications processed yields the matching function

$$m(U, V, \alpha, \gamma, z) = p(z)\frac{aU\gamma V}{aU + \gamma V} \tag{13}$$

The conditional matching function (13) has the usual properties. It is increasing and concave in U, V, α and γ. As V approaches infinity, the contact rate tends to αU and as U tends to infinity, the contact rate is γV in the limit. Similarly, the number of matches approaches zero as either U or V tend to zero. It is also worth noting that the model implicitly captures heterogeneities, as it takes time for workers and firms to search for a good match. If there were no differences, no processing would be necessary. The time-consuming nature of matching is made explicit by α and γ. The time lags in the model are caused by searching, writing applications as well as interviews, for example.

The value of the specification (13) is limited for empirical applications, since α and γ are difficult to observe. However, α and γ can be made endogenous and solved as functions of U and V, leaving the U and V the only arguments in the function.

The interpretation of elasticity estimates obtained by the model deserves more detailed discussion. Elasticities on vacancies and unemployment reflect their search costs relative to search benefits, for workers relative to firms. For example, if the elasticity on unemployment is higher than on firms, it means that the search costs are lower for firms. It follows that firms put greater effort on search and thus impose larger congestion externalities to other firms. (Lahtonen 2006.) Furthermore, in the model the elasticity on unemployment is equal to the probability that an application will lead to a successful match, which corresponds to the proportion of recruiting effort done by firms. Similarly, the elasticity on vacancies equals the probability that a unit of recruitment effort will result in hiring an employee, and again the proportion of search effort undertaken by job seekers. (Hynninen 2007.)

Stevens explains that if the matching function can be modeled as Cobb-Douglas function, then the elasticity with respect to unemployment depends on the bargaining share of workers as well as cost parameters for both firms and job seekers. For instance, higher bargaining share of workers leads to a higher search effort undertaken by the unemployed. This in turn increases the negative congestion externalities that job seekers impose to each other. Similarly, high unemployment benefit has the opposite effect on search effort as it decreases the additional benefit of employment.

As usual, one may question the simplifying assumptions of the model. For example, in reality applications do not completely fail, when a firm is busy, but rather end up in a queue, and job seekers are likely to post several applications simultaneously. However, Stevens defends her model by arguing that relaxing those assumptions would significantly add complexity in the model providing little additional insight. The key feature that the model incorporates is that both parties need to devote some time for assessing the value of a potential match causing congestion effects to other agents.

2.3.3 Mismatch

Petrongolo and Pissarides (2001, p.407) explain the co-existence of vacancies and unemployment as a consequence of aggregation over distinct markets. A good and intuitive starting point for introducing heterogeneities is to think labor market as consisting of several micro markets with limited mobility between those markets. These micro markets may be separated either spatially or in terms of skill or occupation. The important thing is that even though there are no frictions in the micro markets, those micro markets may exhibit imbalance between supply and demand. As the mobility between micro markets is imperfect, at an aggregated level unemployment and vacancies simultaneously exist. In other words, in every micro market the short side clears[7], but adding several micro markets together, vacancies and unemployed add up.

To illustrate the theory let us think of two labor markets A and B that are located distance d from each other and consist of homogenous labor and firms. In A there is not enough demand for labor and thus unemployment whereas in B the economy is growing fast and firms cannot find enough labor to fill the vacancies. In a frictionless economy supply and demand in both regions would converge into equilibrium levels, as unemployed workers in region A would move to region B in response to lack of job in region A and rising wage level in region B. However, there are costs related to moving such as time and effort consumed into finding a new home as well as social costs related to being

[7] $M(U,V) = \min(U,V)$

separated from friends, acquaintances and a possibly relatives. Moreover, there is a possibility that an unemployed worker gets a job in a region A. Therefore for some unemployed it may be optimal to stay in region A and wait. Similarly, jobs cannot easily be transferred geographically since production needs physical capital or demand for services may be elsewhere. As a result, there are simultaneously vacant jobs and unemployed worker if one looks these two labor markets together.

This example gives a very simplified picture of the reality, but illustrates the basic idea. Same principle is easy to transfer in occupational context. Some sectors suffer from lack of skilled labor force whereas in declining industries there are fewer jobs available and hence the unemployment is high. For example, Finland has quite recently experienced a large decline in the number of manufacturing jobs, whereas ever more labor is demanded in healthcare sector, as the population gets older. When taking into account both spatial and occupational dimensions, it is likely that higher level of mismatch is observed. Another way to look at the occupational mismatch is to see it as a mismatch between educational levels and thus skills supplied and demanded.

The theory presented in this chapter, also known as aggregation over distinct markets, has some interesting theoretical implications. As stated by Petrongolo and Pissarides (2001, p.407), if we assume that in each micro market the short side clears and that vacancies and unemployment are exogenously distributed across space, then a CES[8]-type matching function emerges and can be written as

$$M = (U^{-\rho} + V^{-\rho})^{-1/\rho} \tag{14}$$

where $\rho > 0$ refers to the variance of the ratio of unemployment to vacancies across micro markets.

The mismatch hypothesis, yet intuitively appealing, has provided mixed results in empirical studies. Its role has been examined in explaining the observed outward shifts of the Beveridge curve and increased levels of equilibrium unemployment in Europe. The results indicate that mismatch can explain some of shifts of the Beverdige curve and rise in aggregate unemployment, but there are also other shift variables. Manacorda and Petrongolo (1999) study skill mismatch and aggregate unemployment and find that while increased mismatch explains a large proportion of the increased unemployment in Britain, its effect has been negligible in the continental Europe. Albaek and Hansen (2004) study occupational mismatch in the context of Danish labor market and find support for the mismatch hypothesis in accounting for shifts of the Beveridge curve. Bunders (2003) observes that unemployed were increasingly unable to match with new vacancies after mid-90s in Finland. This was due both occupational and geographical reasons.

[8] CES = constant elasticity of substitution

2.3.4 Stock-flow matching

So far we have assumed that firms and job seekers meet randomly. Petrongolo and Pissarides (2001, p.405) suggest that there indeed is a random element in search as there is some luck involved in hearing of vacancies. However, it is likely that systematic search plays a far more important part in the search process. The model presented below assumes that there is perfect information about vacancies and job seekers simultaneously apply for all suitable jobs. It may be that this process well describes the matching taking place through local labor market offices, but finding jobs that are not advertised by LLOs is better depicted as a random process. Gregg and Petrongolo (2005, p.1992) point out that although the reality is likely to be somewhere between the two extremes of the random and stock-flow search, the stock-flow model is attractive in the sense that it incorporate realistic feature of the search process: A job seeker scans lots of vacancy announcements, applies to many and having scanned and rejected an advertisement once, one is less likely to return to it than to a new opening. The model has similarities to the mismatch approach in the sense that a fail to match in the first round stems from mismatch between job seekers and vacancies.

Following Coles and Smith (1998) let us assume that there is complete information about available vacancies and agents can meet at zero cost. Unlike random search models we looked earlier, now job seekers scan the whole market at once and then apply to all jobs that they find likely to be acceptable. Traders negotiate and form a match with some probability. After a period the qualities of all possible matches are revealed and those who have not matched with a trading partner continue to scan new alternatives. Here is key difference to random search model. In random search models negotiation with a trading partner does not change the number of traders on the other side available for a match. Here all scanned opportunities can only be matched by traders who have not scanned them yet. This follows from the assumption that job seekers apply for all suitable jobs and if no match is formed with those firms there is no need to negotiate again. Hence traders who remain unmatched in period 1 can only match with a flow of traders in subsequent periods.

More formally, let there be two types of agents in the market, buyers and sellers. They trade a heterogeneous good, labor input in this case, which has a different value to different buyers. Assume buyers assign a value to sellers´ product that is a draw from some probability distribution and sellers place a zero value to the labor input. To make it simple, let us assume that the value is drawn from a Bernoulli distribution, where the probability that the value $\pi > 0$ is λ and the probability that $\pi < 0$ is 1- λ. These probabilities can be understood as the probability that a match is formed and the probability that no match is formed, respectively. As Lahtonen (2006) points out, a key difference to the urn-ball model and the telephone line model as well is that no frictions between current buyers and sellers exist in the market.

The market operates over discrete time periods of length $\Delta > 0$. Every moment there exists a stock of buyers B_t and sellers S_t respectively. Those stocks consist of agents who have been in the market in the period t - Δ. New agents enter the market following a Poisson process with arrival rates b, s > 0. As Δ tends to zero we obtain the probability that a new buyer immediately matches with a seller

$$P(A\ new\ buyer\ immediately\ matches) = 1 - (1 - \lambda)^{S_t} \tag{15}$$

By dividing (15) by the stock of old sellers we get the probability that a given old seller match with a new buyer.

$$P(A\ particular\ trader\ trades\ with\ a\ new\ buyer) = \frac{1-(1-\lambda)^{S_t}}{S^t} \tag{16}$$

As new buyers enter the market at rate b, we can derive the hazard rate of old sellers

$$h = \frac{b}{S_t}\left(1 - (1 - \lambda)^{S_t}\right) \tag{17}$$

This function is increasing in b and decreasing in S_t. Hence, the model incorporates the usual congestion effects that old sellers impose to each other. Nevertheless, there is no interaction between old buyers and old sellers, which is different from the models presented earlier.

As we assumed a symmetrical behavior of buyers and sellers, similar conditions are valid for a new seller and an old buyer as well. The expected number of matches can be expressed as the sum of matches between new buyers and old sellers and new sellers and old buyers. Thus we obtain a matching function

$$M(S_t, s, B_t, b) = b\left(1 - (1 - \lambda)^{S_t}\right) + s\left(1 - (1 - \lambda)^{B_t}\right) \tag{18}$$

This matching function implies increasing returns to scale. Furthermore, it can be shown that all traders are better of as the flow of new entrants increases. If we look at (16), we see that the transition probabilities of old buyers and sellers conditional on a new potential trading partner on the market are independent of flows s and b. Therefore an increase in the flows clearly makes old traders better off as the entrance interval of new buyers and sellers decreases. Thus the probability of a match in a given time increases.

The model predicts that the transition probability is relatively high after the beginning of an unemployment spell, but if unemployed workers are unlucky in the first round their transition probabilities become significantly lower. This captures the fact that transition probabilities tend to be lower for long-term unemployed. For empirical studies the model indicates that a measure for the duration of unemployment and vacancies is needed in order to account for different trading opportunities for traders that have been in the market for longer time.

In terms of policy implication, Lahtonen (2006, ch.5 p.97) suggests that if the matching process is better depicted as stock-flow process than random search, there is less need for policies aiming at improving the contact process between job seekers and employers. Measures that aim to cohere the preferences and needs of employers and job seekers are more likely to be effective. This may include training job seekers to better fulfill the employer requirements.

The model has also been tested empirically. Coles and Smith (1998) augment their theoretical analysis by running regressions to test the dependencies between stocks and flows of unemployed and vacancies. They find strong support for the stock-flow model, or marketplace model, as they call it. The estimates show that vacancy stocks are insignificant in predicting hazard rates for the unemployed with unemployment durations between 4 and 26 weeks. After 26 weeks of unemployment the stocks are again significant and after 52 weeks the estimated coefficient triple and is highly significant. These effects are, the authors argue, due to start of an ALMP program at 26 weeks and the end of unemployment entitlement at 52 weeks. That makes the unemployed more likely to accept jobs they have turned down earlier. Unlike the stocks, vacancy flows are highly significant with all durations.

Coles and Smith (1998) also test the model with unemployment flows and vacancy stocks. The stock-flow model predicts that stocks of vacancies should be a significant determinant of the unemployment hazard rate for short durations of unemployment. The authors find that unlike for above-month durations, for unemployment durations of 0-1 weeks and 1-2 weeks vacancy stocks are highly significant. They also report that the importance of stocks gradually declines while the role of vacancy flows increases over time. The results also confirm the crowding out effect that unemployed workers cause to each other.

Also Lahtonen (2006, ch.5) finds empirical support for stock-flow model, but there seems to be elements of both random and stock-flow matching processes. The results show that there seems to be little stock-stock matching and the flow of new vacancies significantly increases the hiring probability of job seekers belonging to the unemployment stock. On the other hand, it seems that all job seekers have to wait some time before they match with vacancies. This notion has two different interpretations. In terms of stock-flow model this means that all job seekers need to wait some time before a suitable vacancy emerges. Under random model the interpretation is that there exist frictions in the market that prevent the instantaneous matching, which seems more plausible explanation. For example, Lillrank (2005) finds that it takes a long time before an unemployed job seeker is within employment services causing severe frictions at the beginning of an unemployment spell.

Finally, Soininen (2006) finds evidence for stock-flow matching process, although long-term unemployed have severe difficulties in finding employment in either stock or flow of vacancies. Similarly, Gregg and Petrongolo (2005) find evidence that supports the stock-flow model. The strong support for the stock-flow model has important implications. Since most empirical studies assume random search, the results in the existing literature may be biased if the duration of search by firms and the unemployed is not controlled for.

2.3.5 Ranking

The model presented here has similar conclusions as the stock-flow model in that hazard rates are lower for those with longer unemployment durations, but the reasons are different. In stock-flow framework a longer spell of unemployment decreases the available job opportunities. Here the model assumes that employers treat applications differently whether they are from short-term of long-term unemployed. There are probably several reasons for that. For example, employers may use the duration of unemployment as a proxy for unobservable characteristics or assume decreased level of productivity for those with longer time in unemployment.

Blanchard and Diamond (1994) were first to introduce a model where ranking effects are addressed. The model is an extension to the urn-ball framework discussed earlier. They discuss ranking in terms of the duration of unemployment, but the framework can be extended to also account for other dimensions of heterogeneity, the level of education, for example. As Hynninen (2007, ch.1 p.18) points out, ranking does not affect the total number of matches but determines who will be chosen for a job.

Blanchard and Diamond show that even if all job seekers share the same productivity, ranking model can be applied. Employers may simply use the duration as a hiring rule. Or as well the ranking may stem from above mentioned reasons like decreased productivity or signaling effects. Essential is not why the ranking occurs, but rather what consequences follow.

Empirical literature shows that the share of long-term unemployed is negatively connected to the matching rate (e.g. Hynninen et al. 2009). Although ranking is a mechanism that allocates jobs over individuals and does not directly affect the matching rate, it may still have indirect impact. It may be that ranking prolongs unemployment spells and the prolonged spell of unemployment in turn negatively affects search effort and deteriorates skill level and health leading to a lower overall matching rate. In addition, Anderson and Burgess (2000) find that there is also ranking between unemployed and employed job seekers. It seems that employers prefer employed job seekers over unemployed ones. This is also likely to increase the duration of unemployment.

TABLE 1 Summary of the micro models of the matching process

Model	Source of frictions	Assumptions	Properties and implications
Urn-ball Butters (1977)	- lack of coordination	- homogenous job seekers and vacancies - job seekers apply only for one job at time - firms do not search	- addresses search externalities and time-consuming search - constant returns-to-scale - unrealistic levels and durations of unemployment
Telephone line Stevens (2007)	- lack of coordination - unbalanced search effort between firms and workers	- homogenous job seekers and vacancies - job seekers apply only one job at time - also firms search - firms can process only one application at time	- addresses search externalities and time-consuming search - constant returns-to-scale - elasticity parameters depend on the bargaining shares and search costs of job seekers and vacancies - Cobb-Douglas as a special case
Mismatch Petrongolo & Pissarides (2001)	- imbalance of demand and supply of labor in micro markets (occupational or geographical)	- no frictions in micro markets: short side clears - imperfect mobility between the markets - exogenous distribution of workers and vacancies	- vacancies and unemployment co-exist as the results of aggregation over distinct markets - policies to promote geographical mobility may help - mixed results in empirical literature
Stock-flow Coles & Smith (1998)	-lack of suitable trading partners or inability to match with them	- heterogeneous agents - complete information - job seekers apply to all suitable jobs at once - a job seeker does not return to a vacancy she has already scanned	- after the first period the remained agents can only match with the flow of new traders - re-employment probability is higher in the beginning of the unemployment spell - policies should aim at reconciling the needs of firms and workers, not at improving the contact process - good empirical performance
Ranking Blanchard & Diamond (1994)	-lack of coordination	- extension for urn-ball model - employers rank job seekers over some quality e.g. the length of unem-ployment or education	- hazard rates decrease with the duration of unemployment spell - determines who will get the job, indirect effect on the matching rate

2.3.6 Differences in search activity and reservation wages

The pool of job seekers consists of very different people. As Hynninen (2007, p.16) lists, these individuals can differ in terms of labor market status, age and education. Further sources of heterogeneities are easy to come up with. In addition there are numerous unobservable factors that make individuals different from each other. Common to all these heterogeneities is that they affect the search effort job seekers put in the search and the reservation wage they have when considering job offers. Individual characteristics, both observable and unobservable, make a difference in terms of expected returns and costs related to search, which in turn determine search and reservation wage decisions. The search intensity contributes to the probability of finding a job, whereas reservation wage determines whether an offer is acceptable or not.

Hynninen (2007, p.16) argues that together with heterogeneity in vacancies the different characteristics of jobs seekers need to be addressed in empirical matching function estimations. While it may be justified to treat the matching function as a black box, when it is used as a macroeconomic modeling tool, in empirical studies this does not necessarily give the most relevant results. Hynninen suggests that empirical specifications can address different job seeker groups by including them as shares of the total job seeker pool. These kinds of non-input explanatory variables are referred as shift variables in the matching literature. Some of the shift variables may capture different search intensities and some may approximate different reservation wages. For example, the duration of unemployment spell likely affects the search effort whereas the level of education presumably correlates with reservation wages.

3 EMPIRICAL STUDIES

Matching function has been to a large extent an empirical question. The questions researchers have been interested in have mainly been returns-to-scale, the role of non-input variables, functional specification and recently matching efficiency. Although the studies differ in many aspects, some general conclusions can be made. The studies often find support for Cobb-Douglas specification with approximately constant returns-to-scale, but the results differ depending on the methodology used. Therefore one needs to be cautious when interpreting the results.

The topic is divided in five parts, each discussing a different aspect of the literature. First, methodological issues are discussed. Second the use of different functional forms is deliberated. This is followed by a chapter discussing returns-to-scale and a chapter concentrating on what variables have been found to affect the matching process. Finally, chapter 3.5 surveys research on matching efficiency.

3.1 General methodological issues

Petrongolo and Pissarides (2001, p.393) list four different types of studies, which give evidence on the key matching function idea. First estimates the Beveridge curve type of relation. Lahtonen (2006, p.11) notes that this has the advantage that no flow data is needed and hence stock-flow mismatch problems do no arise. On the other hand an assumption of steady-state relationship between vacancies and unemployment is needed and this indirect way of studying matching has been substituted by direct estimation of the matching function. Second way is to estimate a matching function using data either on whole economy or some sector of the economy, usually manufacturing. Third, many studies use regional data in estimating matching function, which makes sense given regional labor markets are often the most relevant markets for job seekers and firms. The fourth way to study matching is to estimate individual transition probabilities, so called hazard rates. Here only studies directly estimating the matching function are discussed, as those are most relevant given the scope of the study and the empirical application.

The stocks are usually measured at the beginning of a month. Burdett et al. (1993) explain that this is done to avoid simultaneous bias. Simultaneous bias arises when end-of-a-period stocks are used as input variables. The reason is that the flow during a month decreases the stocks and therefore failing to address the problem biases estimates. However, most studies measure the stocks at the beginning of a period and thus avoid the simultaneity problem.

Typically, matching function studies use either monthly or quarterly data. There are also studies using annual data, but this is far more rare[9]. Using

[9] Yearly data are used in Fahr & Sunde (2002) and Ilmakunnas & Pesola (2003), for instance.

annual or quarterly data may cause more severe time aggregation problems. This happens if hires are measured as a flow during a year, whereas input variables are measured as beginning-of-period stocks. Measuring the stocks in the beginning of a period does not necessarily reflect their average value during that month. Burdett et al. (1993) show that if the stocks are mean reverting, there is a downward bias in the following estimates. Mean reversion implies that if the values of stocks are high at the beginning of a period, they tend to decrease over the period and vice versa. If there is a bias its magnitude is proportional to the period, which allows testing the magnitude of the potential bias. The authors find that the bias is not severe when data frequency is monthly. Gregg and Petrongolo (2005) further show how to avoid the time aggregation problem by modifying the matching function specification.

Kangasharju et al. (2005) test how the returns-to-scale in the matching model are affected by the length of measurement period. As predicted by Burdett et al. (1993) the returns-to-scale are significantly higher when monthly data are used instead of quarterly data. The authors also find that using flows of input variables during a period improves the explanatory power of the model by 10 percentage points. Unfortunately, data on unemployment and vacancy flows are often unavailable.

3.2 Variables used

3.2.1 Basic variables

Empirical studies on the matching function differ in terms how the basic variables, namely matches, unemployed job seekers and vacancies are measured. Although different measures can be used, it is important to have such measures that the flow corresponds to the used stocks.

Let us start with discussing different measures for the flow of matches. A commonly used measure is total unemployment outflow. This is however problematic, because it also includes individuals who transit outside the labor force. Moreover, Burgess and Profit (2001, p.316) suggest that the flow out of the labor force is likely to vary in size over the business cycle and across regions. To alleviate the problems arising from omitting the flow out of the labor force, some studies only include unemployed men assuming that those transiting out of labor force are mainly women. If possible, it is advisable to use the outflow from unemployment to employment as the dependent variable when using stocks of unemployment as an input variable, but this is often not feasible due to lack of data. (Broersma & Van Ours 1999, p.78.)

If total hires, vacancy outflow or filled vacancies are used as the dependent variable, the job seeker stock should also include non-unemployed job seekers. For example, Kangasharju et al. (2005) use filled vacancies as the dependent variable, but this is matched with all job seekers as an input variable. Thus, there is correspondence between the dependent and the independent variables. However, using filled vacancies as a measure for matches seems to produce different parameter estimates, but the results are not completely comparable to studies using some form of unemployment outflow. Burgess and Profit (2001, p.316) argue that although the two measures should in theory be the same, they are actually not two noisy measures of the same underlying variable but indeed measure different events.

Studies that measure matches as filled vacancies also include Burgess and Profit (2001) and Broersma & Van Ours (1999)[10]. Those studies find strongly decreasing returns to scale and, moreover, the matching process seems to be driven more by vacancies than unemployment. Broersma & Van Ours test the hypothesis that using filled vacancies while ignoring non-unemployed job seekers leads to downward biased job seeker elasticity. They find that the elasticity with respect to unemployed job seekers is remarkably higher when hires from unemployment are used as the dependent variable instead of the flow of filled vacancies. However, despite the previous evidence, there are also studies measuring matches as filled vacancies and finding constant returns-to-scale (e.g. Van Ours 1991; 1994).

The decreasing returns-to-scale found in Kangasharju et al. (2005) is nevertheless not a surprising result if we think that non-unemployed job seekers have lower search activity than unemployed job seekers. This view is supported by Pissarides (2000, p.123), who argues that employed job seekers have lower search intensity because of higher search costs. Moreover, employed job seekers are also likely to have a higher reservation wage. Thus, it is natural that the return to a pool of all job seekers is lower than in case of a pool consisting only of the unemployed. If only the stocks of unemployment are available, then the outflow from unemployment to employment should be used as the measure for matches (Broersma & Van Ours, 1999).

The variation in results raises the question that how should we measure the matching variables as the results are seemingly dependent on the choice of those variables. Broersma and Van Ours (1999) show that if matches are measured as filled vacancies but non-unemployed job seekers are omitted in the estimation, it results in underestimating the elasticity on job seekers. If the dependent variable is total unemployment outflow, ignoring the stock of non-unemployed job seekers and the flow of non-unemployed to employment biases the results upwards. They argue that the elasticity with respect to vacancies is unbiased and thus the bias in returns-to-scale equals the bias in the job seeker elasticity. The authors test the hypothesis that omitting non-unemployed job seekers biases the results. As expected, they find that including non-unemployed job seekers hardly affects the elasticity on vacancies, but the elasticity with respect to unemployment increases and constant returns-to-scale cannot be rejected anymore.

Moreover, Burgess and Profit (2001, p.316) argue that although the outflow of vacancies and the outflow of unemployment are equal in most theoretical models, in practice they measure different things[11]. The authors find

---

[10] Both fail to account for non-unemployed job seekers, yet the results are similar to those found in Kangasharju et al. (2005).

[11] Also the data in chapter four shows that filled vacancies and unemployment transitions into employment are sometimes very different.

that, in fact, while filled vacancies move pro-cyclically the outflow form unemployment behaves counter-cyclically. The different behavior of flows over the business cycle may explain the different results obtained with different measures of matches. Additionally, Burgess and Profit (2001) find that spillover effects seem to have different effect on filled vacancies and unemployment outflow[12]. To summarize, there is strong evidence that studies having different measures of matches are not comparable and hence, caution is needed when interpreting the results.

The role of on-the-job search deserves some more discussion. In reality there are much more job seekers than there are unemployed individuals. Some people outside labor force[13] search for a job, as do some employed workers. Kangasharju et al. (2005, p.116) explain that besides unemployed workers, a job seeker pool consists of people who are working but willing to switch jobs, workers who are facing a threat of unemployment, people doing household work, workers being temporarily laid off or on a shortened working hours, students and so forth. Their data includes a measure for non-unemployed job seekers and that group is found to comprise almost 40 per cent of all job seekers. Having a slightly different time period Hynninen et al. (2009) find that the share of non-unemployed job seekers is on average 36 per cent of all job seekers. Moreover, Pissarides (2000, p.120) argues that most hires are by workers who are not unemployed. Thus, it seems that employers rank non-unemployed applicants ahead of unemployed ones. Although the importance of non-employed job seekers in matching is undeniable, unfortunately, data on that group is rarely available for empirical studies.

The bias may not, however, be that large, as it first seems. First, when an employed worker matches with a vacancy, the job previously held by this worker now becomes vacant. So we could argue that in this respect employed job seekers do not crowd out unemployed job seekers. On the other hand bias may occur, because now two vacancies are reported although there is only one vacancy from unemployed workers point of view. This means that there may be a downward bias in the elasticity with respect to vacancies if matches of the employed are omitted. Second, Pissarides (2000, p.123) argues that the search intensity of employed job seekers is likely to be lower than the effort undertaken by the unemployed, because of higher search costs. The reason why search costs can be thought to be on average higher for the employed is that if it were otherwise, unemployed job seekers would accept the first offer and then continue search while employed.

Burgess (1993) studies British data and finds that an increase in hiring rate has a much smaller impact on unemployment outflow. In fact the elasticity of hiring rate with respect to unemployment outflow is found to be as low as 0.31, 8 standard errors below unity. Looking at the data, Burgess finds out that when hiring rate increases, the share of new jobs taken by the unemployed decreases. Burgess concludes that this is a result from competition between employed and

---

[12] For example, high unemployment in neighboring regions increases local filled vacancies but reduces local unemployment outflow.

[13] Technically, individuals without a job searching for one should be defined as an unemployed, but in the data this is not always the case.

unemployed job seekers and essentially, the search effort of employed workers is determined endogenously. When more hires take place in the labor market, expected returns for search increase, which makes employed workers to put more effort into search. Also Broersma and Van Ours (1999) and Hynninen et al. (2009) find the number of non-unemployed job seekers to be pro-cyclical. It is noteworthy that using only the unemployed as the measure for job seekers implicitly assumes that the number of unemployed job seekers is proportional to the relevant pool of job seekers (Blanchard & Diamond 1990, p.10).

The contribution of Burgess (1993) is to address the endogenous nature of employed job seekers´ search decision. The implication of the results for empirical matching function studies is the importance to control for the varying number of non-unemployed job seekers even when the outflow from unemployment is used as the dependent variable. Failing to address this issue may bias the returns-to-scale downwards with panel and time series data as high matching rate attracts more job seekers crowding out the unemployed. Anderson and Burgess (2000) join the criticism arguing that commonly estimated matching function parameters are actually reduced-form combination of a matching function we are interested in and a job-competition model like the one in Burgess (1993). Finally, the behavior of non-unemployed job seekers, especially the employed ones, may partially explain why some studies find countercyclical efficiency trends (see e.g. Bunders 2003).

There are also differences in terms how vacancies are measured. While in European studies reported vacancies are often used, many studies with US data use Cenference Board´s help-wanted index as a proxy. Help-wanted index is based on help-wanted advertising in major metropolitan newspapers. Although Warren (1996, p.141) criticizes those series as being “dubious”, Abraham (1987) finds that the index quite closely tracks the movements in the stock of vacancies. Even when there is direct data on vacancies, there are likely to be national differences in terms of what proportion of vacancies are reported in official registers. It is estimated that 50 percent of vacancies in Finland are announced through LLOs (Räisänen 2004). Gregg and Wadsworth (1996) study British data and find that only 30 per cent of vacancies are registered by public labor agencies and only 70 per cent of unemployed use these services as an search channel. However, little is known of how the share of reported vacancies moves over a business cycle. If the share is not constant we run into the same problem discussed above with non-unemployed job seekers. Besides, as Burgess and Profit (2001, p.315) argue, the share may also vary across regions and the vacancy stock likely includes a disproportional share of low skilled jobs, and hence is not a sample of all vacancies.

When regional data are used, it may be important to account for vacancies and job seekers in neighboring regions as well. Intuitively, these vacancies cause a positive externality on the unemployed and a negative one for vacancies in the measurement region. Similarly, job seekers in neighboring areas cause a negative externality on local job seekers and a positive externality on local vacancies. These externalities are generally called spillover effects.

Spillover effects have been tested in the literature[14]. Empirical evidence suggests that spillover effects may play an important role in matching in local labor markets (see e.g. Ilmakunnas & Pesola 2003; Burgess & Profit 2001), although the literature does not give a unanimous view. Burgess and Profit find that after controlling for local labor market conditions, high unemployment in neighboring areas increases filled vacancies, while it decreases local unemployment outflow. This is another reason, why there seems to be different results in the literature depending on what dependent variable is used. More importantly, the authors find that the magnitude of spillover effects varies over the business cycle. When economy is in recession, the unemployed increase the area they search within while firms reduce their own. When economy is booming the effects are reversed. This suggests that if spillover effects cannot be controlled, some measure for the state of an economy may be needed. Burgess and Profit note that most of the previous literature indeed fails to account for spillover effects and the omission of those factors is not rare in more recent studies either. Despite, specific measures of the business cycle are rare in the literature and therefore the stocks of unemployment and vacancies likely capture some of the effects discussed above.

Spillover effects are not only a technical question relating to validity of matching function estimates. If there are spillover effects like what is found by Burgess and Profit (2001), important policy implications follow. It means that local business activity does not only affect neighboring areas by providing jobs, but a lower local unemployment also help other regions within a short distance, by causing less congestion on the job seekers in those areas. Therefore regional policies that aim, for instance, at rescuing failing businesses have positive externalities and thus the costs should be divided between the beneficiaries (regions) of a given policy measure.

Finally, figure 2 summarizes the relationships between the key matching function variables. It can be seen that filled vacancies, unemployment outflow and unemployment outflow into employment all require data that is not regularly available. Using filled vacancies as the dependent variable requires data of non-unemployed job seekers both locally and in neighboring regions as well as information about the number of unemployed job seekers in neighboring regions. When unemployment outflow is used we should have data not only about registered vacancies in neighboring regions (which is sometimes available) but also about the number of non-registered vacancies both locally and in neighboring regions. This illustrates the limitations that imperfect data sets to empirical studies.

[14] A theoretical (and also empirical) analysis is provided by Burda and Profit (1996).

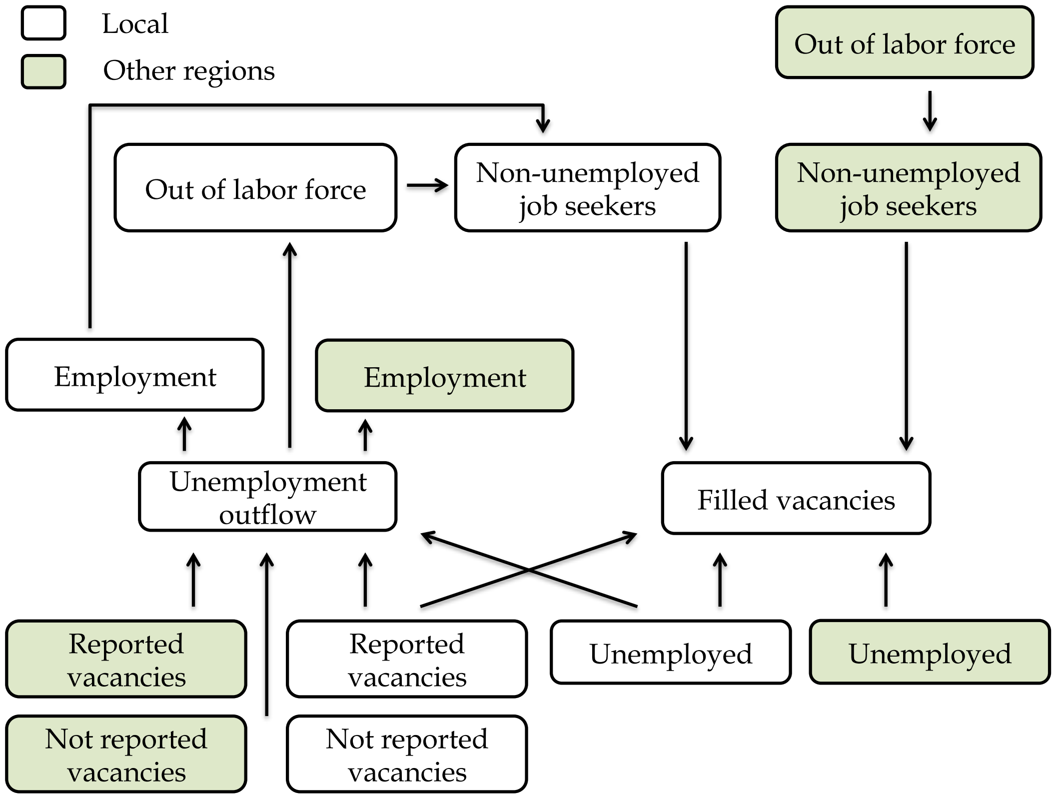


FIGURE 2 Relationships between key variables

3.2.2 Other variables

One important topic in empirical matching function studies has been testing different variables that affect the matching process. The motivation for inclusion of these variables is twofold. First, a shift of the Beveridge curve has been observed suggesting that some additional variables should be included to account for the actual data. Second, micro models suggest inclusion of several structural variables mostly describing the composition of job seeker pool. For example, the stock-flow model indicates that a share of long-term unemployed should be used as an additional variable. To sum up, the data tells us that there are other variables affecting matching and the micro theory and intuition suggest what these variables could be.

The most often included additional variable is the share of long-term unemployed (Lahtonen 2006, p.12). This is also implied by micro models like the ranking model and the stock-flow model. The lower matching rates of the long-term unemployed are also well documented (Gregg and Petrongolo 2005, p.2003). Literature suggests several reasons why the transition rates are lower for the long-term unemployed. First, a long spell of unemployment may decrease the skill level and health, lower the search intensity and have a negative signaling effect for potential employers. Second reason relates to the both observed and unobserved heterogeneity. As individuals with higher productivity are likely to exit unemployment sooner than those with lower level of skills, the pool of long-term unemployed is likely to consist of disproportionately many workers whose hazard rates where below average already when they entered the unemployment pool. Also the stock-flow model has been used as a rationale to include unemployment duration in empirical matching functions.

It is also usual to include a time trend in the matching function. Following Lahtonen (2006, p.12-13) other variables include labor market status (e.g. Mumford & Smith 1999) and flows of new job seekers and vacancies (e.g. Coles & Smith 1998). Also real wages and energy prices (Gross 1997) and GDP growth (Ilmakunnas & Pesola 2003) are sometimes included to approximate the economic situation. Surprisingly, Ilmakunnas and Pesola (2003) find a negative effect of GDP per capita. Finally, some studies include demographics, education levels (e.g. Fahr & Sunde 2002), replacement ratios (Burgess 1993), controls for vacancy quality (Hynninen et al. 2009) and the share of owner-occupied homes (Ilmakunnas & Pesola 2003).

The changes in the matching process over time are likely to be, at least to some extent, labor market specific. However, some general trends are to be found. It seems that a negative trend is more often found than a positive or neutral one. Usually the time trend captures the aggregate change in the output for a given set of inputs. More detailed analysis could further divide the effect into the change in matching technology and the change in efficiency (see e.g. Hynninen et al. 2009) and into the change in the productivity of vacancies and the unemployment stock (see e.g. Gregg & Petrongolo 2005).

Fahr and Sunde (2002) study regional and occupational matching efficiencies using German data. They find that the fraction of the unemployed younger than 25 years old significantly improves matching efficiency while no significant effect is found for the unemployed older than 50 years old. Also Ilmakunnas and Pesola (2003) find a positive relationship between the share of young job seekers and matching efficiency.

A study of Burgess (1993) provides interesting results. Replacement ratio[15] is found to increase unemployment outflow. The author argues that a higher replacement ratio makes the possibility of losing a job less bad and therefore employed workers search less causing less congestion on unemployed job seekers. Although there is likely to be a decreasing effect on the search intensity of the unemployed, the effect of reduced congestion seem to dominate.

Lahtonen (2006, ch.3) studies the effect of education. He finds that the share of low- and high-educated job seekers has a positive relationship to matching productivity, but emphasizes that the effect of education affects through balance of demand and supply of skills in the labor market. Similar results are also found in Fahr and Sunde (2002), yet the effect is much larger for the high-educated. These findings are consistent with the observation that in Finland, as in many other European countries and the US, the share of employment in low-skilled and high-skilled jobs has increased during the past 20 years relative to middle-skill jobs (Goos et al. 2009). Lahtonen suggests that middle-skill workers probably have higher reservation wages than low-skilled

[15] The ratio of unemployment benefit and labor income

workers, which deters them from matching with low-skill jobs and at the same time high-skilled workers are preferred over middle-skill applicants by employers.

Lahtonen (2006, ch.4) also finds support for the hypothesis that densely populated areas are more productive in matching. This hypothesis was presented by Coles and Smith (1996) arguing that in densely populated areas communication requires less effort and is cheaper resulting in higher search activity. On the other hand, Kano and Ohta (2005) suggest that higher population density is also related to higher degree of heterogeneity resulting in an increase in search frictions.

Another interesting factor to study would be the resources of local labor market offices (LLOs), but to my knowledge no study so far has accounted for those issues.

3.3 Functional form

The most used specification in empirical studies is Cobb-Douglas form, which is also referred as log-linear specification (Lahtonen 2006, p.17). The Cobb-Douglas matching function is written as

$$M_t = cU_t^{\alpha}V_t^{\beta} \tag{19}$$

where t denotes time, c is a scale parameter and α and β are elasticities with respect to U and V (see e.g. Lahtonen 2006, p.53). Usually in data U and V are measured at the end of a period and therefore lagged values of those variables are used to obtain the stocks at the beginning of each period. The Cobb-Douglas form is then transformed in log-linear form by taking logarithms. Thus, we obtain:

$$lnM_t = \mu + \alpha lnU_{t-1} + \beta lnV_{t-1} \tag{20}$$

where μ = $\ln c$. Now the elasticity parameters have clear interpretations. α measures the percentage change in matches for one per cent change in the number unemployed. Similarly, β stands for the percentage change in matches for one per cent change in the number of vacancies. (Stock & Watson 2007, p.273.) Hence, we obtain returns-to-scale by adding up the elasticity parameters. If $\alpha + \beta < 1$, there are decreasing returns-to-scale. If $\alpha + \beta = 1$, the matching function exhibits constant returns-to-scale. Finally, if $\alpha + \beta > 1$, the matching technology has increasing returns-to-scale.

Warren (1996, p.135) criticizes Cobb-Douglas specification as being too restrictive. He proposes transcendental logarithmic function, usually called translog function, to be more flexible and therefore minimizing the likelihood of bias arising from misspecifying the underlying matching technology.

There are relatively few studies using translog specification (see e.g. Warren 1996; Yashiv 2000; Fox 2002; Kangasharju et al. 2005). Following Warren (1996) a translog matching function can be written as

$$lnH = \beta_0 + \beta_1 lnU_t + \beta_2 lnV_t + \tfrac{1}{2}\beta_{11} \ln(U_t)^2 + \tfrac{1}{2}\beta_{22}(lnV_t)^2 + \beta_{12}(lnU_t)(lnV_t) + \beta_3 T + \mu_t \, , t = 1, \ldots, N \, , \quad (21)$$

where T denotes a linear time trend and μ_t is an error term following $N(0, \sigma_u^2)$. Interestingly many studies using translog specification find increasing returns-to-scale. The elasticities in this model are obtained as $\epsilon_U = \beta_1 + \beta_{11} lnU + \beta_{12} lnV$ for unemployment and $\epsilon_V = \beta_2 + \beta_{22} lnV + \beta_{12} lnU$ for vacancies. Unlike in Cobb-Douglas, the elasticities obtained from translog model depend on the values of U and V. To calculate numerical values, sample means of U and V are often used (see e.g. Warren 1996; Kangasharju et al. 2005).

Kangasharju et al. (2005) compare Cobb-Douglas and translog specifications using disaggregated data from Finland. Although there is not much difference in the explanatory powers between the models, translog specification gives significantly higher returns-to-scale than Cobb-Douglas. When monthly data including flows of input variables are used the returns-to-scale are 0.99 for Cobb-Douglas and 1.40 for translog specification.

Some studies also use non-linear specifications in estimating a stock-flow model (Gregg & Petrongolo 2005; Lahtonen 2006, ch.5). As there is more support for stock-flow than random search, these studies provide a good benchmark to test whether assuming random search results in biased conclusions about the matching process.

Gregg and Petrongolo (2005) estimate unemployment and vacancy outflow rates taking into account the flows of inputs during a sample interval. When using only vacancy stocks[16] as explanatory variables, the coefficient is significantly positive. When also the inflow of vacancies is included, the estimated coefficient on stocks in no more significant but the parameter on the inflow of vacancies is significantly positive, giving support for the stock-flow model. The same pattern also appears with unemployment stocks and flows when vacancy outflow rate is estimated. Moreover, when looking at elasticity estimates (unemployment outflow as the dependent variable), we find that while the elasticities with respect to the stocks have very small values, the corresponding values on the inflows are about 0.6 for unemployment and 0.35-0.45 for vacancies. When vacancy outflow rate is used as the dependent variable, we find a much higher elasticity estimate for the inflow of vacancies and a much lower estimate for the inflow of unemployed.

The authors also find that the stocks do not match well even with the flow of trading partners suggesting that perhaps those unable to match during the first period are less preferred by the traders on the other side of the market. This does not only relate to the signaling effect of long durations, but also to the true characteristics of traders. For example, vacancies that remain unmatched

[16] There are no stocks of unemployed as the equation is divided by the unemployment stock to obtain the unemployment outflow rate

may be unattractive for job seekers due to low wage or irregular working hours. Similarly, some of those job seekers failing to match may be less preferred by employers due to lack of education and suitable skills. However, the results of further investigation suggest that heterogeneity may not play an important role in explaining the outflow rates of vacancies and the unemployed, although there is reason for caution when interpreting the results.

The implications for the matching literature are following: The matching process seems to be well described by the stock-flow model[17]. Therefore the inflows of the unemployed and vacancies should be incorporated in the empirical matching functions. Moreover, the flows should enter the matching function separately, because of the different impact they have on matching rates.[18] A matching rate of a pool of traders depends on to what extent it consists of stocks and flows. However, as Gregg and Petrongolo use quarterly data, the problem may not be that severe when monthly data are used as also indicated by Burdett et al. (1993). Therefore models using only stocks of the input variables may work as fair approximations with short sample intervals, but nevertheless the stock-flow approach is likely to provide more accurate view of the matching process.

Blanchard and Diamond (1989) also experiment with constant elasticity of substitution (CES) specification. However, later studies have not adopted CES-specification.

3.4 Returns-to-scale and elasticity estimates

Petrongolo and Pissarides (2001, p.397) survey a bulk of empirical literature and find that while some studies find mildly increasing or decreasing returns-to-scale, most of the studies give support for constant returns-to-scale restriction. This means that the size of labor market does not have much effect on matching efficiency. A reason could be that there may be two effects that work to opposite directions. First, in large labor markets there are more opportunities improving the matching process. On the other hand, as there are more job seekers and vacancies, it is likely that the congestion effects are more severe. This in turn has a negative effect on matching efficiency. It may be that those effects compensate for each other and therefore no difference between labor markets of different size is found.

However, one could argue that Petrongolo and Pissarides (2001) overstate the dominance of constant return-to-scale. Kangasharju et al. (2005) note that conclusions about return-to-scale are more difficult to draw when certain

[17] Also Lahtonen (2006 ch.5) finds evidence supporting the stock-flow model, although the matching process seems to exhibit elements of random matching too.

[18] Lahtonen (2006, ch.5 p.98) argues that including the inflows per se would make the results unreliable as the dependent variable depletes the inputs. The second order Taylor approximation of the model in Gregg and Petrongolo (2005) suggests including half of the inflow, which is also suggested by Petrongolo and Pissarides (2001). However, empirically there does not seem to be much difference compared to the full inclusion of the inflows.

extensions to basic Cobb-Douglas model are introduced or when disaggregated data are used. Petrongolo and Pissarides themselves admit that spatial aggregation may bias the estimates towards constant returns-to-scale. Looking at individual studies, there are many cases, which indicate a clear deviation from constant returns-to-scale hypothesis. As discussed earlier, matching function studies differ in many aspects and this together with random variation may cause different results. For example, Broersma and Van Ours (1999) find that returns-to-scale on matching significantly depend on the treatment of non-unemployed job seekers[19]. Therefore one should be careful before making strict conclusions about returns-to-scale. There seems to be no comprehensive summary in the literature studying whether there are clear patterns explaining different results.

The literature is not unanimous whether matching process is driven more by demand or supply of labor and the elasticity estimates are dependent on how the dependent variable is measured. Petrongolo and Pissarides (2001, p.393) argue that when total outflow from unemployment is used, the elasticity on unemployment is estimated to be approximately 0.7 and the elasticity with respect to vacancies about 0.3. When the outflow from unemployment into employment is used, the elasticity on unemployment often slightly drops and is often found to be between 0.5 and 0.7.

Looking at individual studies reveals that there is actually large variation in results. Based on Lahtonen (2006) the range for the elasticities with respect to both vacancies and unemployment is approximately 0-0.9. The large variation in results is probably due to differences in labor markets, different time periods and different variables and methods. Again the results seem not to be general, but rather context-specific and also extra cautiousness should be taken for the methodological issues. It seems that the elasticity with respect to unemployment is higher and the respective coefficient for vacancies lower when total unemployment outflow or the outflow from unemployment into employment is used as the dependent variable. The reasons were already discussed in chapter 3.2.

The reason why the elasticity parameters matter is also that they reveal us the magnitude of externalities that firms and job seekers cause to each other. As discussed before, increasing the number of job seekers decreases the vacancy job seeker ratio meaning lower probability of finding a job. Similarly, increasing the number of vacancies causes a negative externality on existing firms in the market. On the other hand, the externality is positive for the other side of the market as increasing the number of traders to trade with increases the chance of finding a match. Hence, the higher the elasticities are, the lower is the congestion effect with respect to that group of searchers and the higher are the

[19] If the elasticity of non-unemployed job seekers with respect to the number of unemployed job seekers is unity, the authors get decreasing returns-to-scale. When the elasticity is a linear combination of the number of individuals in various non-unemployed groups, an elasticity of 0.1 is obtained resulting in constant returns-to-scale. Finally, when the elasticity is estimated directly within the matching model framework, the elasticity is found to be -0.5 and increasing returns-to-scale cannot be rejected.

positive externalities[20]. In practice this means that, for example, the higher the elasticity on unemployment is, the less the level of unemployment affects individual hiring probabilities (Broersma & Van Ours 1999).

3.5 Matching efficiency

More recently, there has been an increasing interest in studying the efficiency of the matching process. The efficiency is very important, because it partially determines the equilibrium unemployment (Pissarides 2000). This is supported by Hynninen et al. (2009), who find that inefficiencies significantly contribute to the aggregate unemployment rate.[21] Moreover, if we get more detailed knowledge on how frictions affect the matching process, we may be better equipped to design policies to reduce unemployment.

3.5.1 What is efficiency?

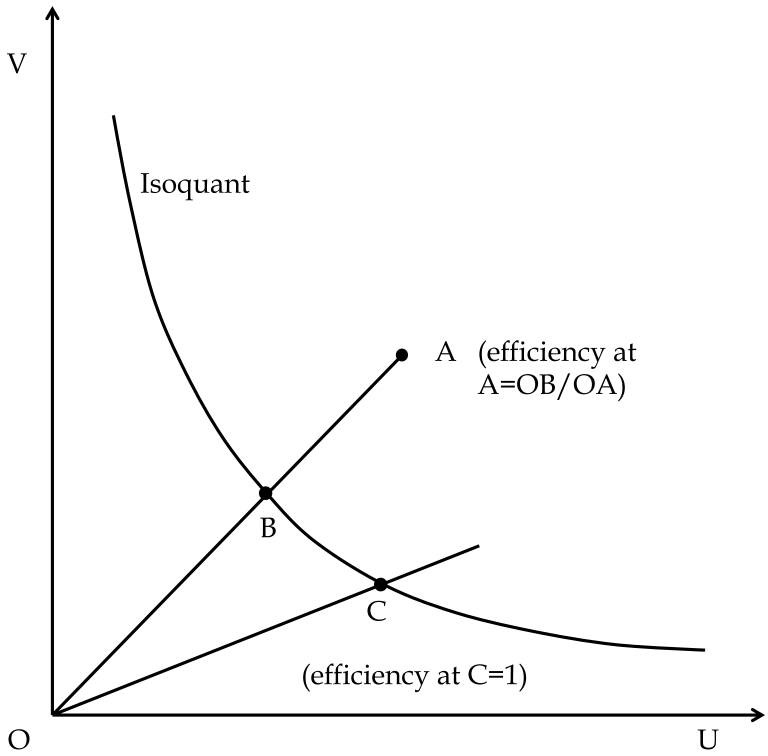
Hynninen (2007, ch.1 p.20) explains that the total efficiency consists of two components. First there is technical efficiency, which measures how many matches are produced at a given set of inputs. The other component is cost efficiency, which in turn relates to proportions of inputs used. The second one is, however, not very applicable in labor market context and therefore this study follows previous literature in using the term "efficiency" as a synonym for technical efficiency.

Figure 3 illustrates the concept graphically. Following Ibourk et al. (2004, p.5) the technical efficiency is measured as the distance between the observed input-output combinations and the technically feasible combinations given by the frontier. In figure 3 the output is fixed and thus we compare the input levels needed to produce that given output. In this context frontier is depicted as an isoquant, which shows the input combinations that can produce a given output if production is technically efficient. Here point C is technically efficient whereas point A is located right from the isoquant and is thus inefficient. Efficiency can be obtained as the ratio OB/OA and inefficiency as 1-OB/OA.

Ibourk et al. (2004, p.5) argue that efficiency is a product of two factors. The first is the rate at which unemployed job seekers and vacancies meet the other being the probability that once met, this contact results in hiring the applicant. The first is determined by search variables like channels used and intensity of search. These are in turn affected by labor market conditions and

[20] Elasticity parameters also bear a close relationship to the question whether a search equilibrium is efficient. For a detailed discussion see e.g. Pissarides (2000, ch.8).

[21] Eliminating the inefficiency not related to observed structural factors would have lowered the aggregate unemployment rate by 2.4 percentage points during the research period. Furthermore, if the unemployed and vacancies in all regions shared the characteristics of those in the most favorable region, the aggregate unemployment would have dropped by 1.4 percentage points.



Source: Ibourk et al. (2004, p.5)

FIGURE 3 The efficiency frontier in U-V space

demographic factors. The latter depends on same factors plus variables that define the quality of a match, e.g. skill requirements.

Gregg and Petrongolo (2005, p.2007) point out that the efficiency can be further divided into two parts. First, how fast vacancies are filled and second, how quickly the unemployed find employment. Usually the two are studied simultaneously, which does not give very accurate and detailed view on the regional or temporal differences in the matching efficiency. The authors use the stock-flow approach and find that while the often-found results of a deteriorating efficiency still applies for unemployment, no such time trend can be found for vacancies. When inter-temporal changes in the efficiency are studied, the disaggregated view may be more appropriate in terms of policy relevance.

Efficiency is also often divided in gross and net efficiencies. As the name suggests, the gross efficiency captures the total deviation from the matching frontier when the structural factors are at their actual levels. Net efficiency in turn sets the structural factors to be equal in all observations. Hence, the net efficiency captures differences in efficiency not explained by the structural variables in the model.

3.5.2 Stochastic frontier analysis

Earlier studies examining the matching efficiencies used a two-stage procedure, where first a fixed effects model was estimated and the regional parameters were then regressed against a set of explanatory variables. As Ibourk et al. (2004, p.7) point out this method not only suffers from the problems common to fixed effects estimator (see chapter 4.2.1 for discussion), but also suffers from a large loss of information and degrees of freedom. Therefore richer concepts are needed.

As the matching function closely resembles a production function, it is natural to apply methods from production theory to estimate the matching function and assess its efficiency. One such method, independently proposed by Aigner, Lovell and Schmidt (1977) and Meeusen and van der Broeck (1977) has been subject to considerable amount of research. This method is called stochastic production frontier analysis.

The key idea in stochastic production frontier models is that a production technology is modeled as a frontier, which measures the maximum technically feasible output at given inputs. Realized outputs are seen as deviations from technically feasible outcomes as depicted in figure 3. This contrasts with traditional panel data methods, where production technology is estimated as an average.

A great feature in stochastic production frontiers is that it enables measuring individual inefficiencies, the deviations from the frontier. Moreover, these inefficiencies can be directly modeled with a set of explanatory variables allowing for a detailed analysis of inefficiencies. Unlike in fixed effect models, the unrealistic assumption of time-invariant inefficiency is often avoided. Instead, inefficiency is assumed to vary over time following changes in both observed and unobserved factors.

3.5.3 Findings about the matching efficiency

Matching efficiency is studied over many dimensions. Usually regional differences are studied (e.g. Ilmakunnas & Pesola 2003; Hynninen et al. 2009; Ibourk et al. 2004). Also occupational labor markets are studied (Fahr & Sunde 2002) and most studies also research the change in the efficiency over time. Below some findings are briefly discussed.

Bunders (2003) finds that matching efficiency is countercyclical. This differs form findings in Ilmakunnas and Pesola (2003) who find a pro-cyclical time trend. This is interesting as both studies study the same period of time. The reason may be that they use different variables in modeling efficiency and more importantly, the studies have different dependent variables. Ilmakunnas and Pesola have a traditional matching function whereas Bunders uses the average duration of a vacancy as dependent variable. This raises the important question how should we measure efficiency in matching. Usually inefficiency is measured as a deviation from either potential or average number of matches during some time period. Although there is a connection between the number of matches and the duration of vacancies and unemployment spells, the concept is still different.

To return to the question of whether efficiency behaves pro- or counter-cyclically, arguments supporting both views can be found. First, we can think that labor market tightness determines how easily job seekers accept offers. In this way of thinking the effect comes from two sources. High unemployment and tight labor market arguably lowers reservation wages and similarly makes job seekers more likely to accept job offers that they would not necessarily accept if the economic situation was better. On the other hand, when finding a job is difficult, it is likely that many job seekers decrease their search effort or some may even stop searching and drop out of labor force.

Most studies on matching efficiencies study regional differences. Bunders (2003) studies the Finnish labor markets and finds that there are large regional differences in the matching efficiency. Analogous to the counter-cyclical movement in the efficiency, Bunders finds that regions with high unemployment rates usually have higher efficiency in matching than ones with lower levels of unemployment. He concludes that high-unemployment regions in Lappland and Kainuu suffer from low demand for labor rather than low efficiency. On the contrary relatively high unemployment rates in Uusimaa and Häme are explained by mismatch between job seekers and vacancies.

Also Ilmakunnas and Pesola (2003) find on somewhat similar results concerning which regions seem to have most efficient labor markets, but there are exceptions and the variance is significantly smaller. The results in Hynninen et al. (2009) do not seem to support the findings in Bunders (2003). As with Ilmakunnas and Pesola, the variance in the gross efficiency scores is small and almost non-existent in the net scores. Again having a different method may explain why the results are different in Bunders (2003) and the other two studies. However, also Ilmakunnas and Pesola and Hynninen et al. have contradictory results. For example, the former finds that Southern Savo has the lowest average efficiency score, whereas in the latter the same region has the highest score. This may be partly explained by different time period, but the conclusion here is that there seems not to be very robust findings concerning regional efficiency differentials in Finland.

As pointed out by Fahr and Sunde (2002, p.17) it is very important to detailed understanding of the nature of the regional unemployment problem at hand in order to design policies to alleviate these problems. Here, efficiency studies may be of great help. In regions or occupations where the matching efficiency is low, policies aiming at increasing the number of vacancies or labor supply may be ineffective. Under these circumstances policy should concentrate more on reducing inefficiencies.

There are nevertheless limitations in the efficiency approach discussed in this chapter. It ignores the quality aspect of matches and only concentrates on the number of matches. The faster vacancies are filled and the faster unemployed job seekers find jobs the better. In fact this may not be optimal. Taking the first acceptable offer may not be desirable, if there is a decent chance that a better offer could arrive soon. This topic goes however beyond the scope of this study. More detailed and formal discussion on the efficiency of the search equilibrium can be found in Pissarides (2000, ch.8).

TABLE 2 Examples of studies estimating the matching function: log-lin, stocastic production frontier and translog specifications

Dependent variable	Authors	Country	Data frequency	Elasticity V	U
log-lin					
FV	Coles & Smith (1996)	England & Wales	monthly	0.3	0.7
FV	Broersma & Van Ours (1999)	Netherlands	monthly	0.5	-0.1
FV	Burgess & Profit (2001)	UK	monthly	0.4	0.003
FV	Kangasharju et al. (2005)	Finland	monthly	0.3	0.1
H	Blanchard & Diamond (1989)	US	monthly	0.6	0.4
H	Anderson & Burgess (2000)	US	quarterly	0.8	0.4
HU	Blanchard & Diamond (1989)	US	monthly	0.2	0.6
HU	Burda & Profit (1996)	Czech	monthly	0.1	0.9
HU	Bleakley & Fuhrer (1997)	US	monthly	0.3	0.6
HU	Broersma & Van Ours (1999)	Netherlands	quarterly	0.01	0.6
HU	Kano & Ohta (2005)	Japan	annual	0.3	0.6
UO	Burda & Wyplosz (1994)	Germany	monthly	0.3	0.7
UO	Burda & Wyplosz (1994)	France	monthly	0.1	0.5
UO	Burda & Wyplosz (1994)	Spain	monthly	0.1	0.1
UO	Burda & Wyplosz (1994)	UK	monthly	0.2	0.7
UO	Mumford & Smith (1999)	Australia	quarterly	0.1	0.9
SPF/log-lin					
H	Fahr & Sunde (2002)	Western Germany	annual	0.35	0.5
HU	Ilmakunnas & Pesola (2003)	Finland	annual	-0.4	0.9
HU	Hynninen et al. (2009)	Finland	monthly	0.04	0.76*
translog					
	Warren (1996)	US	monthly	U+V = 1.3	

Notes: FV = filled vacancies, H=hires, HU=hires from unemployment, UO=unemployment outflow
** True fixed effects model*

4 MATCHING EFFICIENCY – APPLICATION TO FINNICH DATA USING STOCHASTIC PRODUCTION FRONTIER APPROACH

4.1 Data and descriptive statistics

The data used in this study is from the Finnish Social Science Data Archive, which is held by University of Tampere. The data is gathered from registers of local labor offices (LLOs) and spans from 2006 to 2012. It includes variables describing labor force, unemployment, vacancies as well as LLOs themselves. The data is of good quality and has the advantage that it includes transitions from unemployment into employment, which is not always the case in the literature. However, as with most studies on the field also some problems may arise. First, the period includes the financial crisis, which may not be representative for more stable times. The data is on annual level, which is not optimal, but we can alleviate the situation by augmenting the stocks of input variables with respective during-a-period flows. Another problem that arises is that not all jobs are advertised through LLOs. Räisänen (2004) estimates that approximately 50 percent of vacancies are announced in the registers of LLOs. Hence the data does not cover the whole labor market and may not even be a random sample of the total labor market, but rather a selected sample, where low-skill jobs are overrepresented. Moreover, Hämäläinen (2003, p.8) finds that the market share of LLOs seems to be lower in more urbanized regions than in less densely populated areas. But as said, these problems are not unique to this study, but common to most empirical studies in the field.

As offices vary hugely in size, only LLOs having a labor force larger than 30 000 workers are included covering on average 1 893 000 workers annually. Although the cutoff limit is somewhat arbitrary, the idea is to decrease heterogeneities between LLOs. It is likely that the labor markets of the smallest regions function very differently from those with larger labor force. For example, it is plausible that workers in small regions are more dependent on the neighboring regions with respect to job opportunities. This problem could have been avoided if the spillover effects were possible to model, but this difficult because of ambiguous region classification[22].

The variables used to control for different labor force characteristics are motivated by previous studies, although data limitations prevent exact replication. Especially Ilmakunnas and Pesola (2003) and Hynninen et al. (2009) have been of great influence. Yet both studies estimate a stochastic production frontier matching function using Finnish data, the former also uses yearly data enabling reliable comparisons with estimates presented here.

As discussed earlier, there have been many ways to measure both dependent and explanatory variables in the literature. Here the dependent

[22] No information on the geography of the districts was available.

variable, matches, is measured as the logarithm of total outflow from unemployment to employment during a year. The measure is advantageous as it matches with the stock of unemployed, unlike total hires or total outflow from unemployment, which are often used in the literature. Unemployment and vacancies are measured as the logarithms of the total number of unemployed and the total number of vacancies registered at LLOs during a year. Usually stocks at the beginning of a period are used, but in case of yearly data using stocks may result in biased estimates. By using the total number of unemployed workers and vacancies, we can better match dependent and independent variables and thus alleviate the time aggregation problem. This solution is supported by Gregg and Petrongolo (2005), for example, and is also used in Ilmakunnas and Pesola (2003).

When it comes to other variables time index is denoted as *t* and it has values from 0 to 6. The variables that describe the structure of unemployed job seekers and vacancies are as follows: *ShareLTU*, is calculated by dividing the average number of long-term unemployed[23] by the average number of unemployed workers. Similarly, *shareALMP* is calculated by dividing the average number of unemployed in active labor market policies by the average number of unemployed workers. *Share25* and *share55* denote the unemployed workers under 25 years old and over 55 years old as the share of all unemployed respectively. Both measures also include temporarily laid off workers, but better measures were not available. Also a model including *V < 14 days*, a measure for the composition of vacancies in terms how easy they are to fill, is estimated although technical problems cast some doubt on the reliability of the results. *V < 14 days* is measured as the proportion of vacancies filled in less than two weeks out of total vacancies. Unfortunately, a measure for vacancies that are open for a long time was not available.

The variables used here are similar to those used in Ilmakunnas and Pesola (2003), although they have more variables many of which were not feasible here. We have similar measures of U and V, but Ilmakunnas and Pesola use unemployment outflow as a measure for hires. However, the outflow into employment is preferable, because the total outflow from unemployment also contains people transiting out of labor force. Age groups under 25 and over 55 years old are motivated by Hynninen et al. (2009). The idea is that groups consisting of young and old individuals likely have different search intensity and reservation wages and employers may treat them differently from job seekers whose age is somewhere in between. The same argument goes for the share of long-term unemployed, which is also often included in the efficiency studies (e.g. Ilmakunnas & Pesola 2003; Ibourk et al. 2004; Hynninen et al. 2009).

The share of unemployed in active labor market policies is commonly used to control the composition of unemployment stock. As annual data are used, there is no point in using lagged values as in Hynninen et al. (2009) to control the composition change in labor force resulting from the outflow of unemployed with most difficulties in finding employment. Here, as the current period figures are used, the interpretation is that active labor market policies

[23] Individuals who have been unemployed for at least a year

presumably improve the employment prospects of individuals who undergo such policies and it may also proxy the resources of LLOs positively affecting the re-employment probabilities of all unemployed in the region.

A drawback in the study is the lack of direct modeling of spillover effects. As true labor markets may span over borders of individual regions, omitting such effects may bias the results. To understand this, think of a worker living in region A, who finds a job in region B. In the data she is registered as an unemployed in region A and hence also her employment is registered in region A. However the job that she has found is located in region B and not registered in region A. This means that the true number of vacancies is understated in the data. This may result in upward biasing effect on V. However, the effect on V is the opposite for region B, as the vacancy is filled although no match is registered for that region. Since there are mainly relatively large regions in the subset, we can assume the negative bias on V to dominate.

Table 3 below shows some interesting things. Comparing hires from unemployment and the number of filled vacancies we observe that they are sometimes very different. This indicates that there may be spillover effects of such magnitude that may bias the results. However, it needs to be noted that a large share of vacancies is filled by non-unemployed job seekers[24]. Therefore higher number of filled vacancies does not necessarily mean that a region is a net supplier of jobs. To make the interpretation more complicated, unemployed job seekers who find a job through informal channels are not shown in the number of filled vacancies, but are still registered in hires from unemployment. This in turn has an increasing effect on the latter measure relative to the former. Hämäläinen (2003) finds a negative relationship between the degree of urbanization and the market share of LLOs and therefore this effect is likely to be more important in more urbanized regions.

In the capital region there are much more filled vacancies than hires from unemployment. This gives support to our presumption that capital region acts as a large net supplier of jobs. Other big net suppliers of jobs seem to be Tampere and Itä-Uusimaa. On the contrary, regions like Meri-Lappi, Salo, Raisio, Seinäjoki, Kouvola and Pirkanmaa have higher numbers unemployment-employment transitions than total filled vacancies. This indicates that those regions may be more dependent on other regions with respect to job opportunities.

Looking at table 3 we also see that there are still huge differences between regions. Helsinki is ultimately the largest region with a labor force as much as twice as high as in the second largest region Espoo. In terms of the number of vacancies Helsinki has almost one third of all vacancies in the subset. Moreover the capital region comprises almost one third of the total labor force and close to half of all vacancies. To address this issue together with high potential spillover effects, estimations with a subset where Uusimaa region is omitted are also carried out.

[24] Hämäläinen (2003, p.2) finds that 40 per cent of filled vacancies in Finland are job-to-job transitions.

TABLE 3 Key variables by region: 2006-2012 averages

	Labor force	Hires from U (H)	Filled vacancies	U	V
Helsinki	308 864	30 571	118 029	53 399	129 704
Espoo	152 939	6 130	23 831	19 400	25 725
Tampere	115 400	16 165	30 083	29 128	32 320
Oulu region	114 266	26 773	20 484	32 041	22 149
Tikkurila	105 936	5 729	23 091	18 308	25 131
Turku	87 408	17 597	22 550	22 991	25 234
Päijät-Häme	82 813	12 346	12 598	21 220	14 758
Jyväskylä	82 271	19 659	11 681	22 054	12 962
Keski-Pirkanmaa	71 490	12 126	5 845	14 376	6 407
Keski-Uusimaa	69 497	4 240	6 567	9 599	7 219
Etelä-Karjala	62 051	11 593	6 605	16 867	7 322
Joensuu region	53 797	14 003	7 224	17 380	8 090
Kuopio	53 200	10 409	10 950	14 165	12 143
Vaasa	51 715	7 132	10 530	8 789	12 056
Pori region	50 111	10 371	8 672	13 247	9 116
Itä-Uusimaa	45 656	2 621	4 313	7 134	5 209
Kouvola	44 834	10 207	6 217	11 183	6 797
Hämeenlinna	43 825	6 692	8 510	8 865	10 207
Seinäjoki region	43 497	10 724	6 970	9 097	8 098
Pohjois-Uusimaa	42 101	5 202	5 952	6 520	6 655
Länsi-Uusimaa	41 411	3 476	4 416	7 040	5 300
Kotka	40 409	6 749	4 845	10 936	5 344
Raisio	34 800	6 778	3 013	5 633	3 367
Mikkeli	33 426	5 646	4 414	8 244	5 039
Meri-Lappi	30 847	14 117	4 222	10 079	4 767
Salo	30 566	9 141	4 550	6 641	4 965
total	1 893 132	286 197	376 160	404 334	416 083
mean	72 813	11 008	14 468	15 551	16 003
median	52 458	10 289	7 097	12 215	8 094

Although the ratio of vacancies and unemployed workers varies across regions, there are still roughly as many vacancies as unemployed job seekers in total. 90 percent of the vacancies are filled during a year, whereas the number of hires form unemployment is only 70 percent of the number of unemployed workers.

As table 4 above more precisely shows, there is also large variation in hiring rates. Interestingly, hiring rate is above one in Raisio, Salo, Seinäjoki and Meri-Lappi. This is likely to stem from a larger share of temporary jobs. As unemployed are only registered once regardless of the number of unemployment spells, temporary jobs may result in double counting. An unemployed may have more than one transition into employment during a year, even though she is only counted once in the unemployment stock. This may cause an upward bias in unemployment elasticity parameter. Another possible explanation is measurement errors, but given the scale this seems unlikely.

TABLE 4 Descriptive labor market statistics

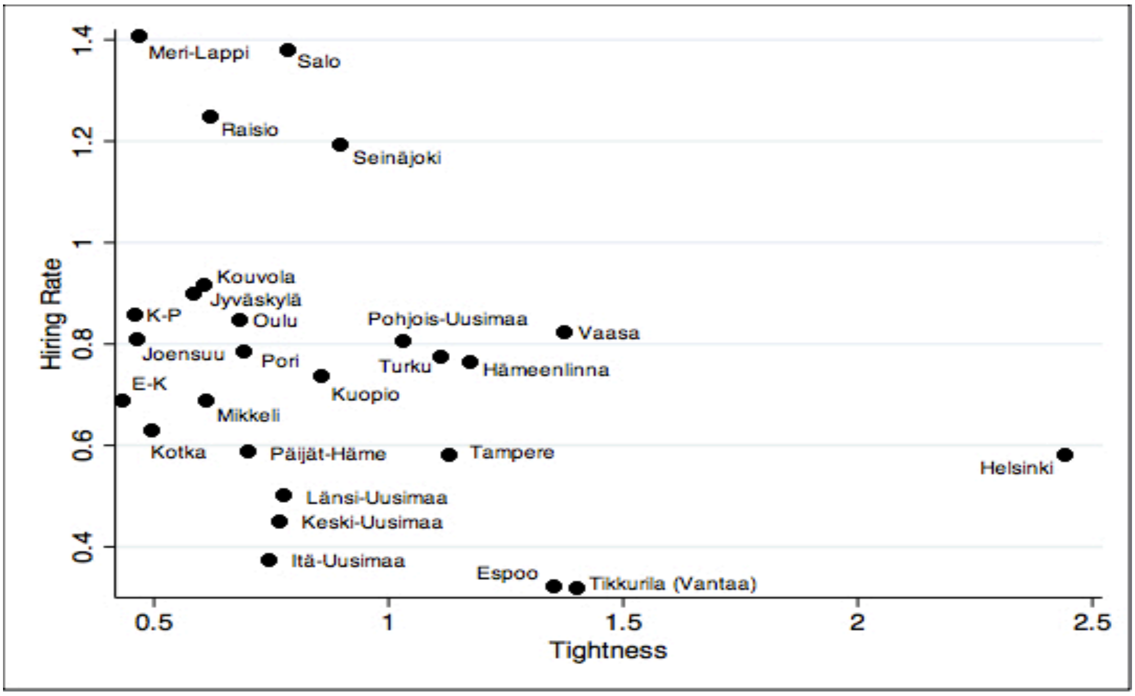
	Unemployment rate (%)	Vacancy rate (%)	Tightness	Hiring rate
Keski-Uusimaa	5.4	10.4	0.77	0.45
Espoo	5.5	16.8	1.36	0.32
Raisio	5.8	9.7	0.63	1.24
Pohjois-Uusimaa	6.1	15.8	1.04	0.80
Vaasa	6.3	23.3	1.38	0.82
Länsi-Uusimaa	6.8	12.8	0.78	0.50
Itä-Uusimaa	6.8	11.4	0.75	0.37
Helsinki	7.5	42.0	2.45	0.58
Tikkurila	7.7	23.7	1.40	0.32
Seinäjoki region	8.2	18.6	0.90	1.19
Keski-Pirkanmaa	8.2	9.0	0.46	0.85
Hämeenlinna	8.4	23.3	1.18	0.76
Salo	9.3	16.3	0.79	1.38
Mikkeli	10.3	15.1	0.61	0.68
Kouvola	10.8	15.1	0.61	0.91
Kuopio	10.9	22.8	0.86	0.73
Turku	11.5	28.9	1.12	0.77
Oulu region	11.7	19.3	0.69	0.84
Päijät-Häme	11.7	17.8	0.71	0.58
Tampere	11.8	28.0	1.13	0.58
Etelä-Karjala	11.8	11.8	0.44	0.69
Jyväskylä	11.9	15.7	0.59	0.90
Pori region	11.9	18.2	0.69	0.78
Kotka	12.4	13.2	0.50	0.63
Meri-Lappi	13.1	15.4	0.47	1.40
Joensuu region	13.9	15.0	0.47	0.81
mean	9.5	18.1	0.88	0.76
median	9.8	16.0	0.8	0.8

Note: Vacancy rate is the ratio of vacancies and the size of labor force. Tightness measures the ratio of vacancies and the unemployed and hiring rate the ratio of unemployment-job transitions and the number of unemployed workers. Figures are averages over years 2006-2012.

Also other variables in table 4 vary a lot across regions. Unemployment rate is more than 2.5 times higher in Joensuu region than in Keski-Uusimaa. Similarly vacancy rate is five time higher in Helsinki than in Keski-Pirkanmaa. Interestingly, all variables seem to be somewhat normally distributed as median and average figures are not far away from each other.

Figure 4 below plots labor market tightness and hiring rates. It is interesting that the relationship seems to be negative. This again gives support for the hypothesis that vacancies and job seekers in neighboring regions play an important role in many labor markets. Moreover, it is also possible that if an unemployed finds a job in another region or a job not announced at a LLO, the transition into employment may not always be registered. This could explain the relatively low hiring rates in Uusimaa.

Two groups stand out in the figure. First, Uusimaa region has very low hiring rates compared to average level. In Itä-Uusimaa, Vantaa and Espoo the hiring rate is less than half of the average level. Helsinki is clearly an outlier in the data with tightness almost twice as high as in Vaasa, the region with the second highest figure. Second, it is easy to see that Meri-Lappi, Salo, Raisio and Seinäjoki have much higher hiring rates than any other region. Explanations for this were discussed above. To address the heterogeneity in the data, the models are also estimated with subsets where both groups are omitted individually and together.



Note: K-P = Keski-Pirkanmaa, E-K = Etelä-Karjala, hiring rate = hires from unemployment / number of unemployed, tightness = vacancies / number of unemployed

FIGURE 4 Hiring rate and labor market tightness

Table 5 below shows the structural variables that are used in modeling the efficiency effects. ShareLTU, shareALMP, share25 and share55 describe the composition of the unemployment pool whereas V < 14 days measures how easy the vacancies in a given region are to fill. It can be seen that although there are differences some of which are large, most of the values vary within a small range. Roughly a third of vacancies are filled during the first two weeks and the job seekers above 55 years old are much more represented in unemployment pools than individuals younger than 25 years old.

Measured by the total contribution to the efficiency, Meri-Lappi has the best structure. However, as there are not many job opportunities, the region suffers from high unemployment. In other words, high efficiency does not necessarily mean low unemployment if there is not much demand for labor. This is supported by Bunders (2003), who finds a positive correlation between unemployment and matching efficiency. A potential explanation could be that high unemployment lowers reservation wages and thus makes people more likely to accept job offers.

Interestingly, the correlation between the share of long-term unemployed and the average share of unemployed in active labor market policies is negative, having a value of -0,5. This is a bit surprising because the long-term unemployed are an essential target of active labor market policies in Finland. On the other hand, it is found that the share of unemployed in active labor market policies seem to move counter-cyclically. (Ilmakunnas et al. 2001.) It is possible that the same pattern also exists across space so that in regions with favorable economic conditions unemployed participate more actively in those programs. It follows that there may also be a reversed causality between the share in active labor marker policies and matching efficiency as is the case with other variables too.

Figure 5 shows how matching rate and vacancy-unemployment ratio have evolved over time. A downward trend in matching rates can be seen during the period. However, the matching rate is much more stable than vacancy-unemployed ratio, which shows large variation. Moreover, matching rate seems to be relatively independent from the vacancy-unemployed ratio. Although the latter exhibits a positive trend in 2006-2008 and 2009-2011, the matching rate has still been decreasing. It seems that during a boom most new vacancies are occupied by non-unemployed job seekers crowding out the unemployed. This is supported by Hynninen et al. (2009), who find that the share of non-unemployed job seekers seem to move pro-cyclically. It is also possible that a high level of job creation changes the composition of a job seeker pool so that those with the best employment opportunities exit the pool first. Thus it follows that the remaining pool may have a lower matching rate despite more vacancies available.

Also a significant drop in vacancy-unemployed ratio can be seen in 2009. However, the matching rate has only followed the trend without a significant drop. The crowding out effect discussed above may work reversely here, as people having a job may be less likely to switch jobs during uncertain times. Also many unemployed with little chance of finding a job are more likely to drop outside the labor force when employment opportunities are weak.

TABLE 5 Structural variables

	shareLTU	Share ALMP	share25	share55	V < 14 days
Helsinki	0,23	0,38	0,08	0,23	0,37
Espoo	0,23	0,34	0,09	0,27	0,31
Tikkurila	0,26	0,33	0,11	0,24	0,35
Pohjois-Uusimaa	0,22	0,46	0,11	0,3	0,29
Keski-Uusimaa	0,24	0,38	0,12	0,3	0,24
Länsi-Uusimaa	0,24	0,39	0,1	0,32	0,27
Itä-Uusimaa	0,27	0,39	0,1	0,3	0,23
Turku	0,26	0,35	0,12	0,23	0,4
Salo	0,22	0,38	0,11	0,29	0,31
Raisio	0,18	0,44	0,12	0,3	0,27
Tampere	0,28	0,33	0,13	0,22	0,42
Keski-Pirkanmaa	0,23	0,42	0,11	0,3	0,32
Kouvola	0,26	0,43	0,13	0,33	0,29
Kotka	0,27	0,38	0,12	0,3	0,25
Etelä-Karjala	0,23	0,44	0,13	0,3	0,33
Mikkeli	0,25	0,46	0,13	0,3	0,28
Vaasa	0,21	0,5	0,14	0,25	0,29
Jyväskylä	0,26	0,41	0,15	0,24	0,34
Kuopio	0,22	0,42	0,14	0,24	0,35
Joensuu region	0,23	0,51	0,13	0,25	0,35
Oulu region	0,22	0,39	0,16	0,18	0,43
Meri-Lappi	0,16	0,45	0,15	0,24	0,33
Pori region	0,28	0,38	0,13	0,31	0,43
Päijät-Häme	0,27	0,36	0,12	0,29	0,27
Hämeenlinna	0,26	0,43	0,13	0,29	0,2
Seinäjoki region	0,22	0,38	0,15	0,23	0,33
mean	0,24	0,41	0,12	0,27	0,32
median	0,24	0,39	0,13	0,29	0,32

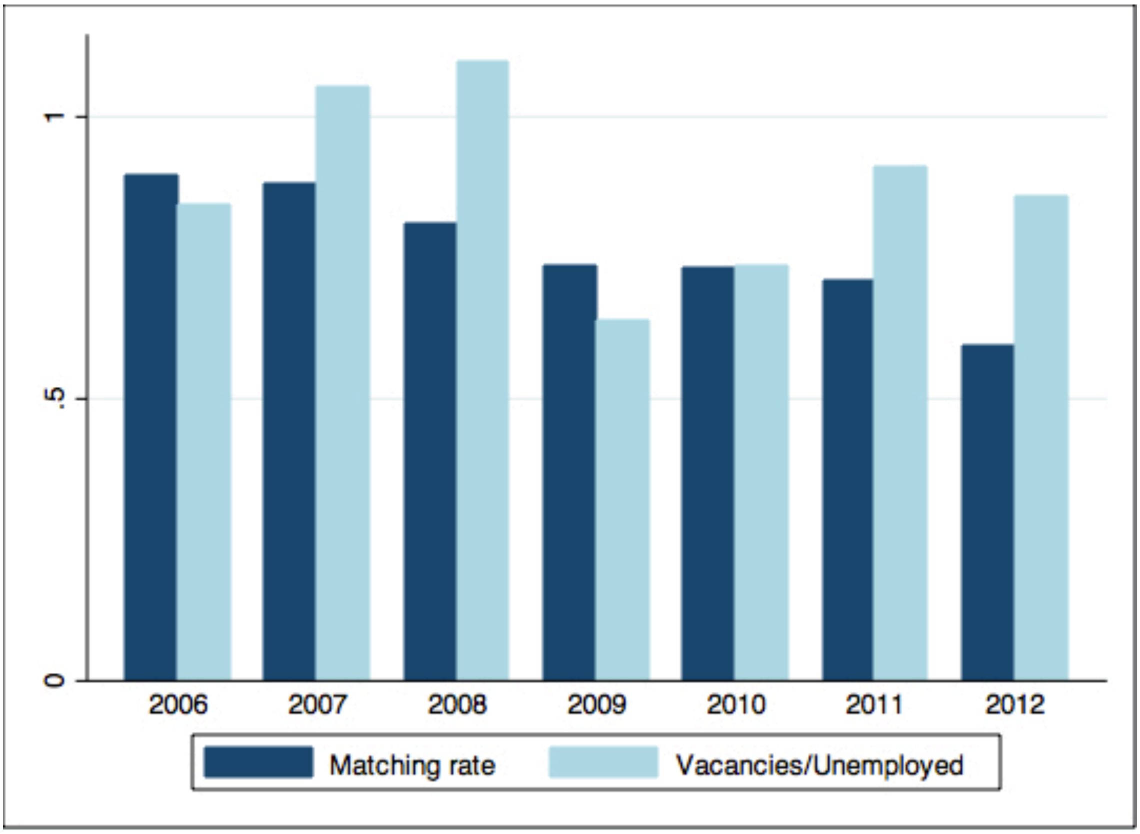


FIGURE 5 Development of matching rate and vacancy-unemployed ratio over time

4.2 Method

4.2.1 Conventional panel data concepts

The matching function is estimated using both conventional panel data methods and stochastic production frontier analysis. The general idea of stochastic production frontiers was discussed earlier in 3.5. Here panel data methods are briefly discussed and then the model is presented.

In this study the conventional panel data methods are used for the sake of comparison. Random effects model, fixed effects model and fixed effects model with suppressed constant term are estimated. The idea behind fixed effects model is that by controlling for region specific unobserved factors that do not change over time, we can obtain more reliable results. Technically this means including region-specific dummy variables and, hence, region-specific intercepts. Fixed effects model has the advantage that the assumptions are less restrictive than random effects model assumptions. On the other hand, fixed effects tend to capture the returns-to-scale effects and even more so if there are large differences in the size of labor market as is the case in this study (Ibourk et al. 2004 p.7).

The use of random effects model is motivated by Hynninen et al. (2009), although the validity of the model is questionable. Random effects model assumes no individual fixed effects, but the individual effects are assumed to be random draws from some distribution. The advantage over fixed effects model is that the random effects model not only uses within-region information, but also between-region information Lahtonen (2006, p.80). The drawback of the method is that it assumes no correlation between random effects and independent variables and if the assumption is violated the random effects estimator is inconsistent. In the context of this study, random effects model does not seem to be theoretically justified, as the observations are not random draws from underlying groups and there is likely to be correlation between explanatory variables and region specific effects. However, the validity of the models can be tested using Hausman specification test, which tests whether the coefficients of two models are statistically significantly different from each other. Under H_0 there is no difference and in that case random effects model is preferable, because it is more efficient. This test gives nevertheless only tentative evidence for the choice of the model, because it only gives valid inference when errors are homoscedastic.

4.2.2 The model

The model used here was developed by Battese and Coelli (1995) and it is widely used in empirical matching function studies with disaggregated panel data (e.g. Ilmakunnas and Pesola 2003; Ibourk et al. 2004; Hynninen et al. 2009). It allows for modeling time varying inefficiency effects by a set of explanatory variables. It is also possible to distinguish the change in technology from time-varying inefficiency. A drawback is that no fixed effects are used in the estimation of the frontier and therefore the model cannot capture unobservable heterogeneities. Another option could have been Greene (2005) true random and fixed effect models, but the comparison with previous literature would have been more difficult as there are not that many studies using the method[25].

In general form the model is

$$Y_{it} = \exp\left(x_{it}\beta + V_{it} - I_{it}\right) \tag{22}$$

where Y_{it} denotes the production, i.e. the flow from unemployment to employment, at region i at time t. x_{it} is a $(1 \times k)$ vector of inputs and other variables defining the production technology with specific values for each region and each time. β is a $(k \times 1)$ vector of unknown coefficients that are to be estimated. V_{it} denotes idiosyncratic error term that is assumed to be independently and identically distributed $N(0, \sigma_V^2)$. That means that the values of V are independently distributed of the I_{it}s. I_{it} can be written as

$$I_{it} = z_{it}\delta + W_{it} \tag{23}$$

[25] To my knowledge, Hynninen et al. (2009) is the only study in labor market context using the method.

I_{it} is a random variable that measures the inefficiency of firm i at time t, thus the time varying inefficiency. By definition, I_{it}s are non-negative and independently distributed following normal distribution truncated at zero with mean $z_{it}\,\delta$ and variance σ^2. z_{it} is a (1 x m) vector of variables explaining the deviations from the frontier and δ is a (m x 1) vector of unknown parameters. The error term of I_{it}, W_{it} is a random variable following a truncated normal distribution with zero mean, variance σ^2 and truncation point at - $z_{it}\,\delta$.

We can obtain the technical efficiency as

$$TE_{it} = \exp(-I_{it}) = \exp(-z_{it}\delta - W_{it}) \tag{24}$$

The technical efficiency is thus the conditional expectation of I_{it} given the above-discussed assumptions.

In this study we define the production technology as Cobb-Douglas form, which is the most used specification in the literature. The vector z_{it} consists of variables presented above. In addition to standard input variables, a time trend is added in x_{it} in some of the specifications. Using the most preferred specification suggested by likelihood ratio test we get

$$\ln H_{it} = \beta_0 + \beta_1 lnU_{it} + \beta_2 lnV_{it} + \beta_3 t + V_{it} - I_{it} \tag{25}$$

where I_{it} is

$$\delta_0 + \delta_1 shareLTU + \delta_2 shareALMP + \delta_3 share25 + \delta_4 share55 + \delta_5 t + W_{it} \tag{26}$$

The model is estimated using Stata and its *sfpanel* command (see Belotti et al. 2012). This command estimates the model using maximum likelihood estimation. Options allow addressing the heteroskedasticity by directly modeling the conditional expectation of σ_V^2 and σ^2 by vectors of explanatory variables. In all cases heterogeneity is observed by finding significant correlations between explanatory variables and σ_V^2 or σ^2 or both. The results are quite sensitive with respect to the choice of those explanatory variables, but when only significant variables are included, the results seem relatively robust.

4.3 Results

First traditional panel data methods were used to estimate the matching function. Panel data methods have the advantage over Battese and Coelli (1995) model in that they are able to differentiate the unobserved time-invariant heterogeneity between regions. The drawback is that inefficiency cannot be differentiated from production technology. In fixed and random effect models the structural variables directly affect the production technology whereas in stochastic frontier models they determine the deviation from the potential output. This profound difference is good to keep in mind when interpreting the results. As discussed above, it is also possible that fixed effects capture returns-to-scale effects causing potential bias on the elasticity estimates.

Table 6 shows the estimation results for fixed and random effects models both for the total sample and a subset where Uusimaa is omitted. The omission of Uusimaa is motivated by figure 4 and the following hypothesis that there are above average spillover effects that may bias the parameter estimates. Indeed, according to Statistics Finland travel-to-work area classification most of Uusimaa region would form a single travel-to-work area, whereas in the data the area is divided into several sub-regions. The same is also true for Turku and Tampere regions. The reason LLOs are not clustered is that it would be difficult to consistently group them even when using the whole data including the smallest LLOs.

The estimation is also carried out using a subset omitting the other cluster to be observed from figure 4, namely Meri-Lappi, Salo, Raisio and Seinäjoki. The model is also experimented with a subset where both clusters are omitted. The results of the preferred specifications, random effects (1) and stochastic frontier model (6) are reported in table 8**.** For comparison, also fixed effect model (3) is reported as the results are slightly different from those of model (1) and as the independence assumption needed to validate random effects is questionable.

The t-statistics in parentheses are calculated using clustered standard errors[26]. The models are compared using Hausman specification test[27]. The test suggests that all models are consistent but random effects model (1) is preferred as it is more efficient than the fixed effects model. However, both (1) and (2) produce very similar results. Also model (3), which is same as model (2) with suppressed constant term produces similar results to (1) and (2) apart from higher and strongly significant elasticity with respect to unemployment. In all specifications the coefficient for *lnU* is highly significant and close to those found in many previous studies using disaggregated data (e.g. Ilmakunnas & Pesola 2003; Ibourk et al. 2004; Hynninen et al. 2009). The coefficient is smaller with models (1) and (2) when Uusimaa region is omitted. Ilmakunnas and Pesola estimate models some of which account for spillover effects and some do not. Interestingly, controlling for spillover effects does not have much impact on unemployment elasticity. The vacancy coefficient in turn becomes remarkably smaller when regional interdependencies are controlled.

All specifications fail to find a significant coefficient on V. Although many studies find relatively small coefficients on V[28], the estimates obtained here are

---

[26] Clustered standard errors allow some misspecification, i.e. heteroscedastic error term, and also over time correlated error structure. Although the consistency of the estimator relies on large samples, Kézdi (2003) shows that it works well even with very small samples. Moreover, Kézdi argues that neglecting autocorrelation is likely to produce larger bias than the small sample bias of the clustered error estimator.

[27] Hausman test cannot be performed with heteroscedasticity-robust standard errors and therefore homoscedasticity-assuming errors are used to test the models. However, the standard errors do not differ much whether robust or non-robust errors are used suggesting the error term is not strongly heteroskedastic.

[28] Ilmakunnas & Pesola (2003) find estimates ranging from 0.2 to 0.4. Ibourk et al. (2004) have coefficients between 0.2 and 0.35. The estimates obtained here are closest to those in Hynninen et al. (2009) that vary between 0.03 and 0.05.

extremely small. The coefficients may be downward biased for the reasons discussed earlier. It is also important to note that as table 3 shows, the vast majority of vacancies are filled during a year. Thus, it seems that local vacancies are often filled by other applicants than local unemployed. In other words, it seems that local unemployed are often crowded out by non-unemployed job seekers as well as job seekers from other regions.

A robust finding across specifications is that there is a negative time trend in matching. This may be due to increased mismatch stemming from higher demands of employers and fast structural change both occupationally and geographically. Another explanation for the finding is that it is possible that the share of matches made through LLOs has changed having two types of effects. First, if proportionately more jobs are announced via LLOs, vacancy productivity decreases when vacancies are measured as jobs announced via LLOs. Second, it is also possible that jobs are more often found through informal channels and these transitions are not always recorded in LLOs.

TABLE 6 Estimation results for fixed (fe) and random effect (re) models

	All (L > 30 000)			**Uusimaa omitted**		
	re (1)	**fe (2)**	**fe (3)**	**re (1)**	**fe (2)**	**fe (3)**
ln U	**0,82**	**0,82**	**0,92**	**0,65**	**0,59**	**0,89**
	(6.79)**	(2.25)*	(12.43)**	(5.78)**	(1.29)	(8.58)**
ln V	**-0,01**	**0,01**	**0,03**	**-0,01**	**-0,01**	**0,06**
	(0.17)	(0.08)	(0.43)	(0.13)	(0.08)	(0.54)
t	**-0,09**	**-0,08**	**-0,08**	**-0,08**	**-0,08**	**-0,08**
	(4.71)**	(3.90)**	(6.35)**	(3.53)**	(3.28)**	(5.16)**
constant	**1,29**	**1,42**		**3,49**	**3,95**	
	(0.94)	(0.33)		(2.76)**	(0.79)	
shareLTU	**0,00**	**-0,13**	**-0,18**	**0,03**	**0,21**	**0,02**
	(0.01)	(0.32)	(0.42)	(0.05)	(0.29)	(0.02)
shareALMP	**0,96**	**0,68**	**0,67**	**0,64**	**0,55**	**0,58**
	(1.89)	(1.32)	(1.96)	(1.13)	(0.88)	(1.34)
share25	**1,74**	**0,05**	**0,35**	**0,35**	**0,60**	**1,74**
	(0.89)	(0.03)	(0.23)	(0.11)	(0.21)	(0.72)
share55	**-0,54**	**-0,74**	**-0,35**	**-1,09**	**-0,82**	**0,44**
	(0.59)	(0.83)	(0.39)	(0.86)	(1.03)	(0.34)

** denotes $p < 0.01$ and * $p < 0.05$

Finally, no significant impact of the structural variables is found. This is different from what is found in SPF-estimations and presented in table 7. The reason may be that unlike Battese and Coelli (1995) model, fixed and random effect models capture the unobserved regional heterogeneities. The results show that when those unobserved factors are controlled for, the effect of structural variables is no more significant. However, this result should be interpreted cautiously and may also be due to lack of data and relatively small variation in structural variables over time, but there is nonetheless evidence that there are likely to be important unobserved factors, which are correlated with explanatory variables causing omitted variable bias. This is not critical as the aim of the study was not to investigate causal relationships between structural variables and inefficiency, but the situation is worth keeping in mind when interpreting the results.

The table 7 below shows the estimates obtained from stochastic frontier models using total sample. The estimates using sample without Uusimaa are not reported here, as they do not differ much from those the total sample. The table of results with Uusimaa omitting subset can be found in Appendix. Following previous studies three different specifications are reported the difference coming from the treatment of time trend. Likelihood ratio test suggests model (6) fits the data best.

The results are close to those obtained using fixed and random effect methods. Also the coefficients with respect to V are higher and significant at 10% level. Moreover, the estimates are very close to those found in Ilmakunnas and Pesola (2003)[29]. Also the elasticities with respect to U are essentially the same as in Hynninen et al. (2009) when using Battese and Coelli (1995) model. Wald test fails to reject constant returns-to-scale hypothesis with models (4) and (6) with p-values 0.57 and 0.41.

Again we find negative time trends. In model (4) the time trend in included in the frontier only. This seems to give a biased view. In model (6) there are time trends both in frontier and inefficiency part and we see that the negative trend over time comes from increased inefficiency not explained by the structural variables. As the inefficiency term enters the model with a negative sign, a positive coefficient indicates higher inefficiency and thus negative impact on efficiency. Ilmakunnas and Pesola (2003) find efficiency to be pro-cyclical. It is possible that the huge decrease in economic activity during the financial crisis[30] causes the negative efficiency time trend of observed magnitude. However, Bunders (2003) finds evidence supporting countercyclical changes over time.

All structural variables are significant and mostly of expected sign. The share of long-term unemployed has a negative impact on matching efficiency. This is likely to be due to factors like ranking, decreased productivity and lower search effort and fewer opportunities. Long-term unemployment also captures mismatch, as those with few suitable jobs in the area are more likely to become long-term unemployed. One may expect this effect to be even larger, but the relative mild effect may be due to the fact that many long-term unemployed drop out of labor force during time. The share of unemployed in active labor market policies enhances efficiency. This probably reflects resources in LLOs as well as improvement of employability of unemployed. Both these findings are in line with predictions as well as earlier literature. The share of young unemployed has a very strong positive effect on efficiency. This effect is also found in other studies, yet the magnitude of the effect is usually found to be lower. The reason behind the observed effect may be a possibly higher search

[29] They find elasticities of 0.9 and 0.1 in stochastic frontier estimations. However, when spillover effects are not controlled, as in this study, the estimates are approximately 0.75 for U and 0.25 for V.

[30] In 2009 the GDP in Finland dropped by 8.5 per cent.

intensity of the young. Young job seekers may also see employment as an investment to future job prospects and therefore are more likely to accept offers. It can be assumed that the group includes disproportionally many job seekers with little education looking for any job they can get and not afraid of taking on temporary jobs. Perhaps surprisingly, the share of job seekers older than 55 years improves the efficiency of matching. A common wisdom goes that those close to retirement are difficult to employ, as employers often see them as more risky hires. A plausible interpretation is that those still in labor force are a selective group consisting of individuals with high motivation to find a job and higher than average employment probability.

TABLE 7 Stochastic frontier estimation: All (L > 30 000)

		(4)	(5)	(6)
Frontier	**ln U**	**0,87**	**0,76**	**0,86**
		(9.30)**	(12.22)**	(9.49)**
	ln V	**0,10**	**0,11**	**0,10**
		(1.82)	(1.89)	(1.76)
	t	**-0,15**		**0,02**
		(15.53)**		(0.78)
	constant	**1,63**	**1,68**	**0,81**
		(3.32)**	(3.96)**	(1.93)
Inefficiency	**shareLTU**	**2,60**	**1,86**	**2,99**
		(4.01)**	(2.82)**	(3.93)**
	shareALMP	**-2,14**	**-6,15**	**-2,27**
		(5.33)**	(6.73)**	(5.69)**
	share25	**-18,93**	**-11,67**	**-18,97**
		(13.24)**	(6.74)**	(13.66)**
	share55	**-5,02**	**-1,54**	**-4,92**
		(10.09)**	(1.77)	(9.36)**
	t		**0,14**	**0,17**
			(14.87)**	(4.88)**
	constant	**5,13**	**3,90**	**4,07**
		(13.94)**	(12.69)**	(12.08)**

Notes: ** denotes p < 0.01 and * p<0.05. In the inefficiency term positive sign means increasing inefficiency and thus lower efficiency. The variance of the technical inefficiency and the variance of the idiosyncratic error term are modeled as functions of shareLTU, shareALMP, share25 and share55.

Table 8 presents the estimation results with subsets where two clusters are omitted in order to test the robustness of the results. First regions with extraordinary high hiring rates seen in figure 4 (cluster 1) are removed from the data. Next, in addition to cluster 1 also Uusimaa regions, which form another cluster in the figure 4, are omitted. As Hausman specification test suggested random effect model to be superior among traditional panel data models, it is also reported here. However, as there is reason to doubt the validity of random effect model assumptions as well as the homoscedasticity assumption needed to perform Hausman test, also fixed effect model (3) is reported. This is preferable specification over fixed effect (2) model, because also previous studies estimate fixed effects models with suppressed constant term. Finally, the stochastic frontier specification that best fits the data as suggested by likelihood ratio test is reported.

First comparing results from models (1) and (3) to those presented in table 6, we see that model (1) gives a higher and model (3) a lower elasticity on unemployment. However, looking at the standard errors (0.12 and 0.07) it seems that the change in results may also be due to random variation. Vacancy estimates are higher than with total sample and the one where Uusimaa region is omitted, although there is no statistical significance when random effects model is used. Structural variables are either insignificant or of the expected sign. Negative and highly significant time component is found in both models. It is also of the same magnitude as what is found in table 6.

Stochastic production frontier model gives significantly higher coefficient on unemployment. The interpretation is nevertheless not straightforward. As the standard error and is very small, it is improbable that the result is a result of random variation. Also testing, for the effect of other possible sets of explanatory variables in modeling the inefficiency error term variance gives support for the high estimate on unemployment. On the other hand, it was difficult to find significant explanatory variables to enable modeling of the variance of idiosyncratic error component possibly leading to biased estimates.

The vacancy coefficient is significant and lower than found in the previous estimation. This can be justified if we assume that unemployed in cluster 1 regions find disproportionally many jobs in neighbor regions. It follows then, that in the remaining sample there are relatively more regions that are net suppliers of jobs. As discussed earlier, this results in underestimating the elasticity with respect to vacancies.

Next both cluster 1 regions and Uusimaa are omitted. The remaining sample is likely to include a lower level of spillover effects and is therefore a good benchmark to test the validity of the results obtained earlier. Although there are still spillover effects that affect the results, the estimates in three last columns indicate that the bias may not be that severe after all. First, it seems that if there is any change in the coefficients, it is likely to be downward, which is in line with predictions. Vacancy coefficient also remains relatively stable, with exception of model (3). Although there may be other explanations to account for the difference, it is again possible that it is due to random variation. On the other hand, same argument can be used to question the validity of other point estimates as well. However, the parameter of model (3) seems to be an outlier and there is much more evidence for vacancy coefficient approximately between 0 and 0.10. Again, the structural variables are either insignificant or of expected sign.

TABLE 8 The preferred models with subsets omitting regions with potentially highest spillover effects

	Regions omitted	Seinäjoki, Raisio, Meri-Lappi, Raisio (Cluster 1)			Cluster 1 and Uusimaa		
		re (1)	fe (3)	SPF (6)	re (1)	fe (3)	SPF (6)
Frontier	**ln U**	**0,99**	**0,84**	**1,09**	**0,82**	**0,71**	**0,78**
(SPF only)		(8.20)**	(12.46)**	(481.22)**	(7.24)**	(6.78)**	(11.55)**
	ln V	**0,06**	**0,13**	**0,01**	**0,12**	**0,23**	**0,12**
		(0.73)	(2.12)*	(3.90)**	(1.36)	(2.74)**	(2.04)*
	t			**-0,09**			**-0,09**
				(212.43)**			(2.02)*
	constant			**-0,66**			**1,33**
				(25.91)**			
Inefficiency	**shareLTU**	**0,07**	**-0,13**	**2,92**	**0,17**	**0,35**	**1,11**
(SPF only)		(0.15)	(0.32)	(3.75)**	(0.20)	(0.50)	(1.74)
	shareALMP	**1,17**	**0,72**	**-3,22**	**0,74**	**0,57**	**-2,39**
		(2.18)*	(2.14)*	(7.11)**	(1.29)	(1.39)	(8.17)**
	share25	**3,51**	**-0,19**	**-14,28**	**2,27**	**2,47**	**-7,49**
		(1.93)	(0.13)	(12.88)**	(0.79)	(0.95)	
	share55	**0,32**	**-1,25**	**-4,18**	**0,31**	**-0,23**	**-1,91**
		(0.31)	(1.77)	(5.76)**	(0.29)	(0.21)	(3.78)**
	t	**-0,10**	**-0,08**	**0,04**	**-0,10**	**-0,09**	**0,03**
		(5.88)**	(6.82)**	(3.75)**	(4.34)**	(5.80)**	(0.67)
	constant	**-1,58**		**3,51**	**-0,17**		**2,33**
		(1.24)		(11.66)**	(0.14)		

Notes: ** denotes $p < 0.01$ and * $p<0.05$. Models (1) and (3) do not separate the frontier and inefficiency, although they are reported under those parts. The signs of the structural variables have different interpretations, so that in (1) and (3) a positive sign means a positive effect and in (6) negative impact. The variance of technical inefficiency is modeled as a function of shareLTU, shareALMP, share25 and share55 and the variance of idiosyncratic error term is assumed constant, as it could not be reliably modeled.

Finally, tables 9 and 10 present estimation results when a measure for vacancy quality, V < 14 days, is used to control for how easy vacancies are to fill in the region. This approach emulates Hynninen et al. (2009) who have controls for jobs filled in 7 and in more than 61 days. Due to lack of suitable data it is only possible to control for the share of easy-to-fill vacancies. The reason why this important aspect is omitted in the baseline specification is purely technical. There is no problem with panel data methods, but with frontier models inability to properly model the heteroscedasticity of the idiosyncratic error term makes the results less reliable.

Fixed and random effects models indicate that controlling for vacancy quality does not have any impact on the results and in particular on vacancy and unemployment coefficients. As can be expected, higher share of easy-to-fill vacancies is associated with a better matching rate.

Table 10 shows the results of stochastic frontier models. Apart from model (3), there is not very much change in unemployment coefficient. However, controlling for vacancy quality strongly increases estimates for the elasticity with respect to vacancies, which is contrary to the expectation. Moreover, the results also show that better quality of vacancies decrease the matching efficiency, which indicates that either there is some sort of omitting variable bias or the model is unreliable. Looking at the results in table 9, it seems that there is no severe omitting variable bias that would justify the counterintuitive sign of V < 14 days and, moreover, controlling for vacancy quality does not have much impact on any coefficients in the model. This makes it tempting to conclude that the inadequate modeling of σ_V^2 results in unreliable estimates. If this is true, also the results in table 8 obtained from frontier models should be treated with doubt.

TABLE 9 Random (re) and fixed effects (fe) models with a control for vacancy quality.

	re (1)	**fe (2)**	**fe (3)**
ln U	**0,81**	**0,80**	**0,92**
	(7.30)**	(2.75)*	(13.36)**
ln V	**-0,04**	**-0,02**	**0,01**
	(0.48)	(0.18)	(0.08)
shareLTU	**0,06**	**-0,05**	**-0,10**
	(0.15)	(0.12)	(0.25)
shareALMP	**1,02**	**0,77**	**0,76**
	(2.09)*	(1.55)	(2.23)*
share25	**1,68**	**0,15**	**0,50**
	(0.98)	(0.09)	(0.35)
share55	**-0,35**	**-0,56**	**-0,10**
	(0.38)	(0.64)	(0.12)
V < 14 days	**0,70**	**0,64**	**0,64**
	(2.85)**	(2.67)*	(3.16)**
t	**-0,08**	**-0,07**	**-0,08**
	(5.12)**	(4.21)**	(6.52)**
constant	**1,32**	**1,58**	
	(1.02)	(0.44)	

Notes: ** denotes p < 0.01 and * p<0.05.

TABLE 10 Stochastic frontier models with a control for vacancy quality.

		(4)	(5)	(6)
Frontier	**ln U**	**0,81**	**0,81**	**0,61**
		(8.37)**	(8.29)**	(6.40)**
	ln V	**0,25**	**0,25**	**0,34**
		(3.59)**	(3.49)**	(5.17)**
	t	**-0,15**		**0,08**
		(12.75)**		(6.90)**
	constant	**0,94**	**0,14**	**0,92**
		(1.81)	(0.29)	(1.95)
Inefficiency	**shareLTU**	**1,01**	**1,14**	**0,87**
		(1.76)	(1.90)	(1.61)
	shareALMP	**-2,70**	**-2,81**	**-2,20**
		(7.39)**	(7.27)**	(5.87)**
	share25	**-18,98**	**-19,09**	**-21,87**
		(10.49)**	(10.56)**	(12.43)**
	share55	**-5,83**	**-5,62**	**-5,84**
		(15.24)**	(12.83)**	(14.49)**
	V < 14 days	**0,85**	**0,85**	**0,89**
		(5.01)**	(4.75)**	(4.95)**
	t		**0,15**	**0,24**
	constant	**5,87**	**5,00**	**4,99**
		(14.33)**	(14.27)**	(16.69)**

Notes: ** denotes $p < 0.01$ and * $p<0.05$. In the inefficiency term positive sign means increasing inefficiency and thus lower efficiency. The variance of technical inefficiency is modeled as a function of shareLTU, shareALMP, share25 and share55 and the variance of idiosyncratic error term is assumed constant, as it could not be reliably modeled.

5 CONCLUSIONS

Modern labor markets are characterized by dynamics that reallocate labor from less profitable work to areas where the return is higher. Jobs are destructed and workers laid off but at the same time new jobs are created and new employees hired. Labor markets are nonetheless imperfect and therefore the adjustment does not happen immediately. There are many sources of frictions that prevent the instantaneous matching of job seekers and vacancies. For example, it takes time to get within unemployment services, screen for suitable jobs, apply and negotiate for the mutual suitability of employment relationship.

This complicate process has been modeled in the literature using the aggregate matching function, which captures key features of the matching process without imposing intractable complexity into economic modeling. The concept has attracted substantial attention in both theoretical and empirical literature. Theoretical research on the topic has mainly engaged in establishing micro foundations for justifying the use of certain functional forms, returns-to-scale and non-input explanatory variables. Empirical literature in turn has estimated the matching function using both aggregated and disaggregated data and testing for different functional forms and explanatory variables as well as studied intertemporal behavior of the matching process. Recently, more emphasis is put on modeling the efficiency of the matching process.

The conventional wisdom in the literature seems to be that there exists a matching function that is well approximated by a Cobb-Douglas function with constant returns-to-scale. However, there are several weaknesses in the matching literature especially when it comes to the empirical modeling of the matching process. First, data are often imperfect and different measures of the variables seem to produce different estimates. Due to lack of data spillover effects, non-unemployed job seekers and non-registered vacancies are often omitted potentially biasing the results as discussed in chapter 3. Besides, the bias is likely to amplify if the composition of job seekers and vacancy pools changes over the business cycle as the literature suggests. Moreover, the results are sensitive to functional specification. While more restrictive Cobb-Douglas specification has a tendency to often exhibit constant returns-to-scale, increasing returns-to-scale are often found with more flexible translog specification. Here the theory cannot help us either to judge between different approximations of an unknown technology. Finally, most studies implicitly assume random matching process, despite evidence in favor of the stock-flow approach.

The survey was followed by an empirical analysis with recent data taking advantage of stochastic frontier methods. The data is comparable to what is used in the previous literature, but like many previous studies, also estimations carried out in this thesis suffer from some methodological flaws. The data does not allow the inclusion of non-unemployed job seekers and non-registered vacancies, which may bias the results if the number of all vacancies and job seekers behaves differently from the registered ones, either over time or space. The study also omits spillover effects, but the potential bias is tested by estimating the model without regions with presumably highest interaction with neighboring areas suggesting that the bias is likely to be moderate. Also the measurement interval is relatively long and monthly data would have been more preferable. Annual data imprecisely measures how fast matches are formed and thus hide an essential element of frictional labor markets. Finally, Battese and Coelli (1995) model does not have any panel effects and therefore fails to differentiate regional fixed effects from inefficiency.

The empirical study appends the recent line of studies analyzing matching efficiency contribution coming mainly from the recent data. The results, mostly in line with previous studies, give tentative support for the constant returns-to-scale hypothesis. The elasticity with respect to unemployment mainly varies between 0.7 and 0.9, whereas vacancies seem to have less impact on unemployment-employment transition and the elasticity is mostly between 0 and 0.10. Furthermore, these results are relatively robust across specifications and different subsets of the data. Nevertheless, due to likely disproportionate variation in the share of registered vacancies over time and space, the vacancy coefficients should be interpreted to concern only the registered vacancies. In this respect the external validity of the study is limited.

Although there is some variation in the point estimates and the vacancy estimates are likely to be slightly downward-biased for several reasons discussed earlier, we can reliably conclude that the number of job seekers is likely to be the most important explaining factor of unemployment-employment transitions, at least when measured at an annual basis. Unemployed job seekers often find a job during a year, but this process does not seem to be helped much by the number of local registered vacancies because of crowding out effect caused by other job seekers. There is evidence that employers often prefer employed job seekers over the unemployed ones making the return to employment more difficult for the unemployed (Anderson & Burgess, 2000).

It is tempting to interpret the elasticities as causal parameters. This follows not only from the economic theory and common reasoning, but also the fact that regional differences can be controlled with explanatory variables and panel data methods. A counterargument for the high matching productivity of the unemployed could be that regions and years with good employment prospects may show more unemployed relative to the non-unemployed pool, because more people join the labor force if employment opportunities are good.

If we interpret the results in causal terms, it has the implication that measures to promote labor supply in extensive margin, i.e. creating incentives for individuals outside the labor force to participate in the labor market, is likely to be more effective in increasing local employment than attempts to enhance job creation. Creating vacancies may also be effective but it does not seem to help much the local unemployed. If policy aims to alleviate local unemployment problem by creating more jobs, those jobs should perhaps be targeted to match with the qualities possessed by the local unemployed. Creating similar jobs that are already in the market does not seem to significantly increase the local matching rate. On the other hand, making predictions based on the model is not that simple, because in reality the number of vacancies and unemployed are driven by factors which also affect the number of non-registered vacancies and the behavior of non-unemployed job seekers, which may remain unaltered if a policy to promote labor supply or to create more vacancies is established.

The coefficients on structural variables are robust in sign. Moreover, in terms of sign these are quite similar to those found in previous research. It seems that long-term unemployment is negatively related to matching efficiency, whereas the share of young and old unemployed is positively correlated with the efficiency. This is especially true for the young job seekers. Similarly, the share of unemployed in active labor market policies is found to have a positive relationship with matching efficiency possibly due to improved skills of the unemployed and also potentially signaling the resources of local labor offices. In all cases, it is likely that a reverse causality also exists. Together with the limited number of explanatory variables, it makes the coefficients purely descriptive.

Finally, a very robust finding across all specifications is a negative time trend. This development, if the evidence is correct, would be a fruitful source of future research. A potential explanation could be the fast structural change that has been taking place in the Finnish economy. Another challenge for future research would be to acquire data that allows modeling all relevant factors of the matching process.

REFERENCES

Abraham, G. (1987) Help-Wanted Advertising, Job Vacancies, and Unemployment. Brookings Papers on Economic Activity 18(1), 207-248.

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

Albaek, K. & Hansen, H. (2004) The Rise in Danish Unemployment : Reallocation or Mismatch? Oxford Bulletin of Economics and Statistics 4(66), 515–536.

Anderson, P. M. & Burgess, S. M. (2000) Empirical matching functions : estimation and interpretation using state-level data. The Review of Economics and Statistic 82(1), 93–102.

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

Belotti, F., Daidone, S., Ilardi, G. & Atella, V. (2012) Stochastic frontier analysis using Stata. The Stata Journal, 1–39.

Blanchard, J. O. & Diamond, P. (1989) The Beveridge Curve. Brookings Papers on Economic Activity, 1-76.

Blanchard, O. J. & Diamond, P. (1990) The Aggregate Matching Function. NBER Working Paper 3175.

Blanchard, O. J. & Diamond, P. (1994) Ranking, Duration , Unemployment and Wages. The Review of Economic Studies 61(3), 417–434.

Bleakley, H. & Fuhrer, J. C. (1997) Shifts in the beveridge curve , job matching, and labor market dynamics. New England Economic Review, 3-19.

Broersma, L. & Van Ours, J. C. (1999) Job searchers, job matches and the elasticity of matching. Labour Economics 6(1), 77–93.

Burda, M. C. & Profit, S. (1996) Matching across space: Evidence on mobility in the Czech Republic. Labour Economics 3(3), 255–278.

Burdett, K., Coles, M. & van Ours, J. (1994) Temporal aggregation bias in stock-flow models. CEPR Discussion Paper 967.

Burgess, S. & Profit, S. (2001) Externalities in the matching of workers and firms in Britain. Labour Economics 8(3), 313–333.

Burgess, S. M. (1993) A model of competition between unemployed and employed job searchers : an application to the unemployment outflow rate in Britain. The Economic Journal 103(420), 1190–1204.

Coles, M. G. & Smith, E. (1996) Cross-Section Estimation of the Matching Function : Evidence from England and Wales. Economica 63(252), 589–597.

Coles, M. G. & Smith, E. (1998) Marketplaces and Matching. International Economic Review 39(1), 239–254.

Davis, S.J. (2005) Comment on ‘‘Job loss, job finding, and unemployment in the U.S. economy over the past fifty years’’. In: Gertler, M., Rogoff, K. (Eds.), NBER Macroeconomics Annual. MIT Press, Cambridge, 139–157.

Diamond, P. A. & Maskin, E. (1979) An equilibrium analysis of search and breach of contract, I: steady states. Bell Journal of Economics 10, 282-316.
Diamond, P. A. (1982) Wage determination and efficiency in search equilibrium. Review of Economic Studies 49, 217-27.
Fahr, R. & Sunde, U. (2002) Estimations of Occupational and Regional Matching Efficiencies Using Stochastic Production Frontier Models. IZA Discussion Paper 552.
Goos, M., Manning, A. & Salomons, A. (2009) Job polarization in Europe. The American Economic Review 99(2), 58–63.
Greene, W. (2005a) Fixed and Random Effects in Stochastic Frontier. 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(2), 269–303.
Gregg, P. & Wadsworth, J. (1996) How effective are state employment agencies? Jobcentre use and job matching in Britain. Oxford Bulletin of Economics and Statistics 58, 443–467.
Gregg, P. & Petrongolo, B. (2005) Stock-flow Matching and the Performance of the Labor Market. European Economic Review 49(8), 1987-2011.
Gross, D.M. (1997) Aggregate Job Matching and Returns to Scale in Germany. Economic Letters 56(2), 243-48.
Hall, R. E. (1979) A theory of the natural unemployment rate and the duration of employment. Journal of Monetary Economics 5, 153-69.
Hall, R. E. (2005) Job Loss, Job Finding, and Unemployment in the U.S. Economy over the Past Fifty Years. NBER Working Paper 11678.
Hämäläinen, H. (2003): Työvoiman rekrytointi toimipaikoissa vuonna (2002). Ministry of Labour, Helsinki.
Hosios, A. J. (1990) On The Efficiency of of Matching and Related Models of Search and Unemployment. The Review of Economic Studies 57(2), 279–298.
Hynninen, S. (2007) Matching in Local Labour Markets Empirical Studies from Finland. Jyväskylä Studies in Business and Economics 56.
Hynninen, S.-M., Kangasharju, A. & Pehkonen, J. (2009) Matching Inefficiencies, Regional Disparities and Unemployment. Labour, 23(3) 481–506.
Ibourk, A., Maillard, B., Perelman, S. & Sneessens, H. R. (2004) Aggregate Matching Efficiency□: A Stochastic Production Frontier Approach , France 1990 – 1994. Empirica 31, 1–25.
Ilmakunnas, P. & Pesola, H. (2003) Regional Labour Market Matching Functions and Efficiency Analysis. Labour 17(3), 413–437.
Ilmakunnas, S., Romppanen, A. & Tuomala, J. (2001) Työvoimapoliittisten toimenpiteiden vaikuttavuudesta ja ennakoinnista. Vatt research reports 77.
Kano, S. & Ohta, M. (2005) Estimating a matching function and regional matching efficiencies: Japanese panel data for 1973–1999. Japan and the World Economy 17(1), 25–41.
Kézdi, G. (2003) Robust Standard Error Estimation In Fixed-Effects Panel Models. Hungarian Statistican Review Special 9, 96–116.

Lahtonen, J. (2006) Matching Heterogeneous Job Seekers and Vacancies - Empirical Studies Using Finnish Data. Jyväskylä Studies in Business and Economics 50.

Manacorda, M. & Petrongolo, B. (1999) Skill Mismatch and Unemployment in OECD Countries. Economica 66(262), 181–207.

Meeusen, W. & Broeck, J. Van Den. (1977) Efficiency estimation from Cobb-Douglas production function with composed errors. International Economic Review 18(2), 435–444.

Ministry of Labour and the Economy: Tutkimus ja ennakointiryhmä (2013). Työ- ja elinkeinotoimistojen kokonaisaineisto 2006 - 2012 [electronic data]. FSD2853, version 1.0 (2013-08-08). Finnish Social Science Data Archive, Tampere.

Mortensen, D. T. (1982) The matching process as a non-cooperative bargaining game. In The Economics of Information and Uncertainty (ed. J. McCall). NBER University of Chicago Press, Chicago.

Mortensen, D. T. & Pissarides, C. A. (1994) Job creation and job destruction in the theory of unemployment. The Review of Economic Studies 61(3), 397–415.

Mumford, K. & Smith, P.N. (1999) The Hiring Function Reconsidered: on Closing the Circle. Oxford Bulletin of Economics and Statistics 61(3), 343-364

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

Pissarides, C. A. (1984) Efficient job rejection. Economic Journal (Supplement: Conference Papers) 94, 97-108.

Pissarides, C. A. (1985) Short run equilibrium dynamics of unemployment, vacancies and real wages. American Economic Review 75, 676-90.

Pissarides, C. A. (2000) Equilibrium unemployment theory. MIT Press, Cambridge.

Rogerson, R., Shimer, R. & Wright, R. (2005) Search-Theoretic Models of the Labor Market: A Survey. Journal of Economic Literature 43(4), 959–988.

Soininen, H. (2006) Empirical studies on labor market matching. Publications of the Swedish School of Economics and Business Administration 159.

Stevens, M. (2007) New Microfoundations for the Aggregate Matching Function. International Economic Review, 48(3), 847–868.

Stock, J. & Watson, M. (2007) Introduction to Econometrics. Second edition. Pearson Education, Boston.

Warren, R. S. (1996) Returns to scale in a matching model of the labour market. Economic Letters 50, 135–42.

Yashiv, E. (2007) Labor search and matching in macroeconomics. European Economic Review 51(8), 1859–1895.

APPENDIX

TABLE 11 Summary of the independent and dependent variables by year

Variable	Year	Obs	Mean	Std.Dev.	Min	Max
Flow into employment	2006	26	12 861	8 456	2 742	33 911
	2007	26	11 716	7 677	2 440	31 426
	2008	26	10 598	7 029	2 298	29 630
	2009	26	11 455	6 891	2 939	32 601
	2010	26	11 324	6 765	2 930	31 470
	2011	26	10 290	6 387	2 623	29 559
	2012	26	8 809	5 459	2 374	25 397
	total	182	11 008	6 991	2 298	33 911
Unemployed during a year	2006	26	15 086	10 284	4 864	52 985
	2007	26	14 040	9 586	4 513	48 667
	2008	26	13 813	9 280	4 254	46 230
	2009	26	16 995	11 115	6 130	56 834
	2010	26	16 902	11 207	6 971	57 093
	2011	26	15 877	10 864	6 513	55 163
	2012	26	16 148	11 219	6 185	56 818
	total	182	15 551	10 426	4 254	57 093
Vacancies during a year	2006	26	14 592	20 280	3 585	105 969
	2007	26	17 149	25 576	3 776	134 447
	2008	26	17 382	25 854	3 560	136 902
	2009	26	13 009	19 812	2 841	105 342
	2010	26	14 814	22 171	3 236	117 420
	2011	26	17 857	29 808	3 173	157 588
	2012	26	17 219	28 575	3 399	150 259
	total	182	16 003	24 496	2 841	157 588
Share of long-term unemployed	2006	26	0,27	0,03	0,21	0,34
	2007	26	0,26	0,03	0,18	0,31
	2008	26	0,23	0,03	0,16	0,29
	2009	26	0,18	0,03	0,09	0,23
	2010	26	0,22	0,04	0,13	0,29
	2011	26	0,25	0,04	0,15	0,32
	2012	26	0,26	0,04	0,17	0,32
	total	182	0,24	0,05	0,09	0,34

Variable	Year	Obs	Mean	Std.Dev.	Min	Max
Share of unemployed in active labor market policies	2006	26	0,34	0,05	0,25	0,44
	2007	26	0,41	0,05	0,31	0,51
	2008	26	0,42	0,05	0,33	0,51
	2009	26	0,35	0,05	0,26	0,45
	2010	26	0,40	0,07	0,28	0,57
	2011	26	0,46	0,06	0,35	0,60
	2012	26	0,45	0,05	0,35	0,56
	total	182	0,40	0,07	0,25	0,60
Share of unemployed under 25 years old	2006	26	0,11	0,02	0,07	0,16
	2007	26	0,10	0,02	0,06	0,15
	2008	26	0,11	0,02	0,07	0,16
	2009	26	0,14	0,02	0,09	0,17
	2010	26	0,13	0,02	0,09	0,17
	2011	26	0,13	0,02	0,08	0,16
	2012	26	0,13	0,02	0,09	0,17
	total	182	0,12	0,02	0,06	0,17
Share of unemployed over 55 years old	2006	26	0,28	0,05	0,19	0,36
	2007	26	0,30	0,05	0,20	0,38
	2008	26	0,29	0,05	0,19	0,36
	2009	26	0,23	0,03	0,16	0,30
	2010	26	0,24	0,03	0,17	0,31
	2011	26	0,27	0,03	0,19	0,34
	2012	26	0,27	0,04	0,19	0,34
	total	182	0,27	0,04	0,16	0,38
V < 14 days	2006	26	0,32	0,07	0,22	0,50
	2007	26	0,31	0,07	0,20	0,44
	2008	26	0,31	0,07	0,19	0,47
	2009	26	0,33	0,09	0,16	0,55
	2010	26	0,35	0,09	0,13	0,58
	2011	26	0,31	0,07	0,12	0,43
	2012	26	0,29	0,08	0,15	0,43
	total	182	0,32	0,08	0,12	0,58

TABLE 12 Summary of describing variables by year

Variable		Obs	Mean	Std.Dev.	Min	Max
Unemployment rate (%)	2006	26	9,5	2,8	5,1	14,4
	2007	26	8,5	2,7	4,4	13,3
	2008	26	7,9	2,7	3,9	13,0
	2009	26	10,3	2,8	6,0	15,4
	2010	26	10,3	2,5	6,1	13,9
	2011	26	9,7	2,6	5,4	13,4
	2012	26	10,0	2,8	5,8	14,2
	total	182	9,5	2,8	3,9	15,4
Vacancy rate (%)	2006	26	17,3	7,0	9,6	35,6
	2007	26	19,9	8,2	11,1	44,6
	2008	26	20,1	7,6	10,3	45,1
	2009	26	14,8	5,9	7,1	34,1
	2010	26	16,7	6,6	7,3	37,4
	2011	26	19,3	8,6	8,3	50,0
	2012	26	18,4	8,5	7,5	47,6
	total	182	18,1	7,6	7,1	50,0
Tightness	2006	26	0,84	0,39	0,38	2,00
	2007	26	1,05	0,52	0,49	2,76
	2008	26	1,10	0,55	0,47	2,96
	2009	26	0,64	0,32	0,30	1,85
	2010	26	0,74	0,37	0,32	2,06
	2011	26	0,91	0,51	0,41	2,86
	2012	26	0,86	0,49	0,36	2,64
	total	182	0,88	0,47	0,30	2,96
Hiring rate	2006	26	0,89	0,33	0,36	1,59
	2007	26	0,88	0,32	0,34	1,59
	2008	26	0,81	0,30	0,31	1,44
	2009	26	0,73	0,29	0,31	1,39
	2010	26	0,73	0,31	0,30	1,63
	2011	26	0,71	0,33	0,29	1,62
	2012	26	0,59	0,27	0,26	1,45
	total	182	0,76	0,32	0,26	1,63

TABLE 13 Stochastic frontier models with Uusimaa omitted.

		(4)	(5)	(6)
Frontier	**ln U**	**0,76**	**0,77**	**0,77**
		(14.56)**	(13.92)**	(14.55)**
	ln V	**0,12**	**0,11**	**0,11**
		(3.44)**	(3.34)**	(3.40)**
	t	**-0,12**		**-0,06**
		(20.06)**		(2.71)**
	constant	**1,71**	**1,48**	**1,46**
		(4.88)**	(3.85)**	(3.97)**
Inefficiency	**shareLTU**	**-0,18**	**-0,17**	**-0,25**
		(0.32)	(0.29)	(0.44)
	shareALMP	**-1,76**	**-1,77**	**-1,79**
		(8.36)**	(8.07)**	(8.46)**
	share25	**-9,02**	**-9,02**	**-9,10**
		(9.49)**	(9.63)**	(9.59)**
	share55	**-0,08**	**-0,11**	**0,03**
		(0.12)	(0.18)	(0.05)
	t		**0,12**	**0,06**
			(20.24)**	(2.54)*
	constant	**2,31**	**2,10**	**2,10**
		(11.33)**	(11.90)**	(11.48)**

Notes: ** denotes $p < 0.01$ and * $p < 0.05$. In the inefficiency term positive sign means increasing inefficiency and thus lower efficiency. The variance of the inefficiency term and the variance of the idiosyncratic error term are modeled as functions of shareLTU, shareALMP, share25, share55.