UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Master Degree Project in Economics

The Matching Process and Regional Disparities

An Efficiency Analysis of the Swedish Labour Market

Linnea Glans and Nezanet Habtom

Supervisor: Conny Wollbrant
Graduate School

Abstract

This thesis investigates the matching process between unemployed and unfilled vacancies in Swedish regional labour markets. Using a stochastic production frontier analysis, we are able to study differences in matching efficiency between regions and changes in overall matching efficiency over time. We use monthly labour market data from the Swedish Public Employment Service for time period 2006 to 2016. Our results suggest that more densely populated counties, on average, have lower matching efficiency than less densely populated counties. We also find that matching efficiency has deteriorated during the research period. Further, the deterioration is robust across all 21 counties. This is consistent with the public belief about a persistent decline in Swedish labour market efficiency after the financial crisis 2008. However, there is no visual evidence of an increase in the negative trend in conjunction with the crisis. This indicates that the deterioration in matching efficiency probably started prior to the crisis.

Keywords: Matching efficiency, stochastic frontier, regional labour markets, financial crisis, Sweden

Contents

1 Introduction

In the wake of the financial crisis in 2008, the global economy faced a process of rapid deterioration in terms of GDP declines and unemployment growths. The economic recession and the following labour market downturns led to structural imbalances on the labour market that have proven to be difficult to reverse during the economic recovery (Lindvall, 2012; Mossfeldt and Österholm, 2011; OECD, 2014). Even though Sweden recovered more rapidly than other countries and often is referred to as a success story, some challenges and structural problems remain (Anxo and Ericson, 2015; OECD, 2015).

Following the crisis, policy makers and experts have raised concerns about deteriorating matching efficiency - the ability of the labour market to match unemployed workers to available jobs - in the Swedish labour market. While unemployment in Sweden remains relatively high, companies in various industries are reporting labour shortage. Declining matching efficiency and persistent unemployment could be a result of increasing skill, sectoral and regional mismatches on the Swedish labour market (Håkansson, 2014; OECD, 2014; 2017).

Labour markets are often studied from a national perspective, where one assumes that labour market efficiency is the same in all regions. In the beginning of 2017, the Swedish Public Employment Service reported declining unemployment rates in most Swedish counties. The Swedish Public Employment Service shows, at the same time, that there are considerable differences between regional unemployment rates (Arbetsförmedlingen, 2017). Given that regional differences exist on the Swedish labour market, it is important to complement the aggregated studies with disaggregated studies in order to implement adequate policies (Aranki and Löf, 2008; Arbetsförmedlingen, 2017). Therefore, the purpose of our thesis is to study the Swedish matching process from a regional perspective. The thesis aims to investigate differences in regional matching efficiency as well as changes in efficiency that may have taken place in the aftermath of the financial crisis.

The economic literature provides various frameworks to study labour market efficiency and the relationship between unemployed and vacancies, but the most commonly used approach is the matching function. The matching function was developed in resistance to the frictionless neoclassical model of the labour market and has become an important tool in the field of empirical macroeconomics and labour economics due to its simplicity (Mortensen and

Pissarides, 1994; Petrongolo and Pissarides, 2001). The model is based on the idea that the labour market is characterized by various frictions, which prevent the market to immediately match job seekers with unfilled vacancies (Petrongolo and Pissarides, 2001). Some frictions are inevitable as the search and matching process is time-consuming. But if frictions mainly are a result of inefficiencies in the matching process, then frictions may cause higher structural unemployment than necessary (Hynninen, 2009).

In order to empirically estimate regional inefficiencies in the matching process, we employ a stochastic production frontier analysis. The stochastic frontier model treats the matching function as a production function and inefficiencies are measured as deviations from optimal output given the inputs. The data used in this study is derived from the Swedish Public Employment Service (SPES) and comprises monthly labour market data for the period 2006:1-2016:12. The data contains registered job seekers and vacant jobs reported to the local labour offices (LLOs). As matching efficiency may depend on labour market heterogeneities, we also include variables representing the composition of the unemployment stock according to age, labour market status and education.

This thesis contributes to the empirical literature in two ways. To start with, according to our knowledge, we are the first to study labour market inefficiencies in Swedish regions using a stochastic frontier approach. The stochastic frontier approach is advantageous to conventional panel data methods in the sense that matching inefficiency is allowed to vary over time and can be separated from time-invariant heterogeneity (Hynninen, 2009; Ibourk et al., 2004). Secondly, our data set contains the years before and after the financial crisis in 2008. Analyses of the Swedish matching process in regional labour markets are rare and with several years apart. The novelty of our data contributes to literature by providing a picture of how Swedish regional labour markets have developed in the aftermath of the crisis. This could be of interest both for Swedish policy makers who have to react on future labour market shocks and for other countries with labour market conditions similar to those in Sweden.

Our empirical analysis suggests that there are distinct differences in matching efficiency between Swedish counties. Stockholm has, on average, the lowest matching efficiency during the time period. Gotland, on the other hand, is the county with the highest. Overall, the results indicate that more densely populated counties have lower efficiency estimates than less densely populated counties. Our results also indicate a pronounced negative trend in median matching efficiency during the research period. Furthermore, the negative trend is consistent across all regions. Namely, all Swedish counties have lower efficiency estimates in 2016 than in 2006, even though some counties have seen larger changes than others. These results are consistent with the belief of deteriorated matching efficiency in the Swedish labour market in the aftermath of the financial crisis in 2008. However, our results suggest no visual structural break in the negative trend in efficiency during or after the years of the crisis, indicating that the deterioration in efficiency started even before the financial crisis broke out.

Our thesis is closely related to the study of Aranki and Löf (2008), which investigates Swedish regional differences in matching efficiency during 1992-2007 using conventional panel data methods. In accordance with our result, they find significant differences in efficiency among Swedish regional labour markets and that matching efficiency tends to be lower in counties that are more densely populated.

This thesis is also related to the matching literature based on the stochastic frontier approach. Hynninen et al. (2009) and Hynninen (2009) use this method studying regional labour markets in Finland. Their results indicate that regional inefficiencies are an important factor of total unemployment. Similar to our thesis, they also include the composition of the unemployment stock in their matching function and find that it has a significant impact on the matching efficiency. Ilmakunnas and Pesola (2003) conduct a similar study also using Finnish regional data. They find that business cycle fluctuations have a large impact on regional matching efficiency.

The rest of this thesis is organized as follows: Section 2 describes the concept of the matching function and introduces the stochastic frontier approach. Section 3 contains a literature review, which is divided into two parts. The first part focuses on studies of labour market mismatches in the Swedish context. The second part discusses empirical research on labour market efficiency based on the stochastic frontier method. Section 4 includes a description of our data and some descriptive statistics. Section 5 describes the empirical method used to study the matching process and inefficiencies in Swedish counties. Section 6 discusses the results of the regional matching process and the efficiency analysis. Finally, section 7 draws conclusions, considers policy implications and discusses further research.

2 Theoretical Framework

2.1 Matching Functions

2.1.1 Random Matching and Stock-flow Matching

The matching function describes the process of job creation as a function of workers searching for jobs and firms searching for workers to fill their vacancies. The model is based on the idea that various frictions characterize the labour market, which prevents the labour market from clearing instantaneously. Frictions are often seen as a result of information imperfections, heterogeneities, mobility problems and other similar factors. As a consequence, the transition from one job to another is often characterized by a spell of unemployment before a new job is found. The effectiveness in matching unemployed job seekers to available vacancies is determined by the intensity of frictions. The stronger frictions, the less effective is the labour market and the matching rate is reduced. In order to overcome these issues, both sides of the labour market usually have to make some effort in terms of search activity (Forslund and Johansson, 2007; Petrongolo and Pissarides, 2001).

The matching function was developed in resistance to the frictionless neoclassical model of the labour market where the important aspects of these issues are ignored (Mortensen and Pissarides, 1994). The model has therefore become a widely used modelling tool to investigate how frictions influence equilibrium outcomes in the labour market. The most commonly used functional form for the matching function is the Cobb–Douglas form (Kano and Ohta, 2005; Petrongolo and Pissarides, 2001):

$$M = m(U, V) = AU^{\alpha}V^{\beta} \tag{1}$$

where M denotes the number of jobs formed during a given time interval, U represents the number of unemployed individuals searching for jobs, and V represents the number of unfilled vacancies. A represents a scale parameter, which depends on various factors that promote or hinder the matching among job seekers and firms. Furthermore, α and β can be interpreted as elasticities with respect to the size of the unemployment and vacancy pools.

The matching function is assumed to be strictly concave and increasing in both its arguments. This implies that matches are more likely when the number of unemployed or vacancies increases, but with diminishing marginal returns in matching. The function is also usually assumed to exhibit constant returns to scale (Petrongolo and Pissarides, 2001). Testing for constant returns of scale has played an important role in the empirical literature and the returns to scale seem to depend on factors such as measurement frequency and level of disaggregation of the data. Among the literature, evidences of both increasing (e.g. Anderson and Burgess, 2000; Boeri and Burda, 1996 etc.) and decreasing returns of scale (e.g. Aranki and Löf, 2008; Hynninen, 2009 etc.) have been found in contrast to constant returns to scale.

The mechanism of labour market matching is often described by two different technologies: random matching and stock-flow matching. In the random matching model, which is described by equation (1), the number of matches is assumed to only depend on the pools of unemployed and vacancies in the beginning of each time period. The pool of unemployed is treated as homogenous and assumed to apply for vacant jobs by random (Blanchard and Diamond, 1994; Petrongolo and Pissarides, 2001).

In the stock-flow model, on the other hand, the pool of unemployed is treated as heterogeneous with perfect information about the labour market. Moreover, the number of matches is assumed to depend on the within-period inflow of job seekers and vacancies in addition to the stocks. The perfect information assumption is based on the fact that job seekers apply for jobs by first scanning a lot of advertisement rather than by pure randomness. Also, it is less likely that a job seeker returns to the same job advertisement than looking for new vacancies once an offer has been rejected (Coles and Smith, 1998; Gregg and Petrongolo, 2005; Petrongolo and Pissarides, 2001). Based on the stock-flow matching, the random Cobb-Douglas matching function can be rewritten as:

$$M = m\left(U, V, \dot{U}, \dot{V}\right) = AU^{\alpha_1}V^{\beta_1}\dot{U}^{\alpha_2}\dot{V}^{\beta_2} \tag{2}$$

The main difference between random matching, equation (1), and its counterpart stock-flow matching, equation (2), is the inclusion of the within-period inflow of job seekers, $\dot{U}$, and vacancies, $\dot{V}$. Thus, the elasticities is now in regard of both stocks and inflows of unemployed job seekers and vacancies, respectively (Aranki and Löf, 2008).

According to stock-flow matching, some of the job seekers flowing into unemployment will find matches within the period, while some will not. The inflow of job seekers who do not find a match immediately, will eventually match with the new inflow of vacancies and leave unemployment (Dmitrijeva and Hazans, 2007).

Against this background, inefficiencies in the random matching model can be explained by coordination failure. Both unemployed and firms are not sufficiently well informed to find suitable matches. On the other hand, in stock-flow matching, inefficiencies are a result of heterogeneities. Mismatches occur since there are no suitable matches between unemployed and vacancies. Therefore, the choice between random and stock-flow matching has an impact on how potential inefficiencies in the labour market could be explained (Aranki and Löf, 2008).

2.1.2 Search Intensity, Reservation Wage and Ranking of Job Seekers

The simplest form of the matching function does not take into account that unemployment durations may be influenced by worker heterogeneity. However, empirically it has been found that individual characteristics are important factors explaining differences in unemployment durations and in the probability of finding a job. Worker heterogeneity can be introduced in the matching function by assuming that individual characteristics have an impact on the search intensity. Since search effort is assumed to represent a cost in terms of time and other resources, every individual chooses the amount of search effort that maximize the net return. For this reason, individuals will choose different amounts of search efforts, which will be based on their search cost, unemployment cost, and the expected returns of employment. The aggregate matching function including search intensity can be written as (Petrongolo and Pissarides, 2001):

$$M = m(sU, V) \tag{3}$$

where s is the average number of search units supplied by the unemployed and sU is the total number of search units supplied. This implies that an unemployed individual that supplies search units finds a job with the probability $s_i m(sU, V)/sU$. Furthermore, the average transition rate for unemployed job seekers can be expressed as $m(sU, V)/U$.

Petrongolo and Pissarides (2001) argue that this justifies the inclusion of aggregated variables that influence individual decisions during search when estimating matching functions. This could for instance be demographic variables that capture differences in search intensity such as the youth share of the population, if it is assumed that young people potentially search with lower intensity than adults. It could also be variables that influence the cost of search, for instance unemployment insurance.

Another way to include worker heterogeneity in the matching function is to assume that individual characteristics affect the reservation wage. The individual chooses a reservation wage and rejects all offers below the reservation wage. If the worker or the firm rejects the offered wage, a match will not be formed. As the function with search intensity, seen in equation (3), this extension of the matching function justifies the inclusion of aggregated variables that could affect the reservation wage. This could for instance be the proportion of unemployed with different education (Petrongolo and Pissarides, 2001).

A third alternative way to include worker heterogeneity in the matching context is to study the search process from the view of the employers. This is called ranking and implies that employers treat applications differently depending on individual characteristics. Blanchard and Diamond (1994) present a model where unemployed are divided into two groups, short-term and long-term unemployed. The model assumes that employers always rank short-term unemployed over long-term unemployed. The duration of unemployment could for instance be seen as a proxy for unobservable characteristics related to ability or productivity. The implication of ranking by unemployment durations is that short-term unemployed only compete against other short-term unemployed in the hiring process. Whereas the long-term unemployed competes with both groups, but only gets jobs where there are no applications from short-term unemployed. The matching function with ranking can be written as:

$$M = m(U^S + U^L, V) \tag{4}$$

where U^S represents the number of short-term unemployed and U^L represents the number of long-term unemployed. The market is assumed to suffer from a coordination failure so that the long-term unemployed are unaware of which jobs the short-term unemployed apply for. This implies that the transition rate from unemployment into employment for short-term unemployed is $m^S(U^S, V)/U^S$, whereas the rate for long-term unemployed is $m(U^S + U^L, V)/U^L - m^S(U^S, V)/U^S$ (Petrongolo and Pissarides, 2001).

2.2 The Stochastic Production Frontier

The stochastic frontier approach was preliminary used in theoretical analyses of firms' behaviour, but has then been applied in empirical studies of the matching function in labour markets. In the stochastic frontier literature, the matching function is treated as a production function that determines the optimal output (i.e. the number of successful matches) that could be produced by the given number of unemployed job seekers and vacant jobs. The inefficiencies are measured as deviations from the optimal output for a given level of inputs. By utilizing the stochastic production frontier, one can estimate the efficiency of the matching process and examine how it varies over time and regions. Moreover, the ability to produce successful matches given available inputs can be modelled using the concept of technical efficiency. The technical efficiency is defined as the ratio of the optimal level of production and the actual production. The greater deviation, the lower is the technical efficiency (Fahr and Sunde, 2002; Ilmakunnas and Pesola, 2003; Ibourk et al., 2004).

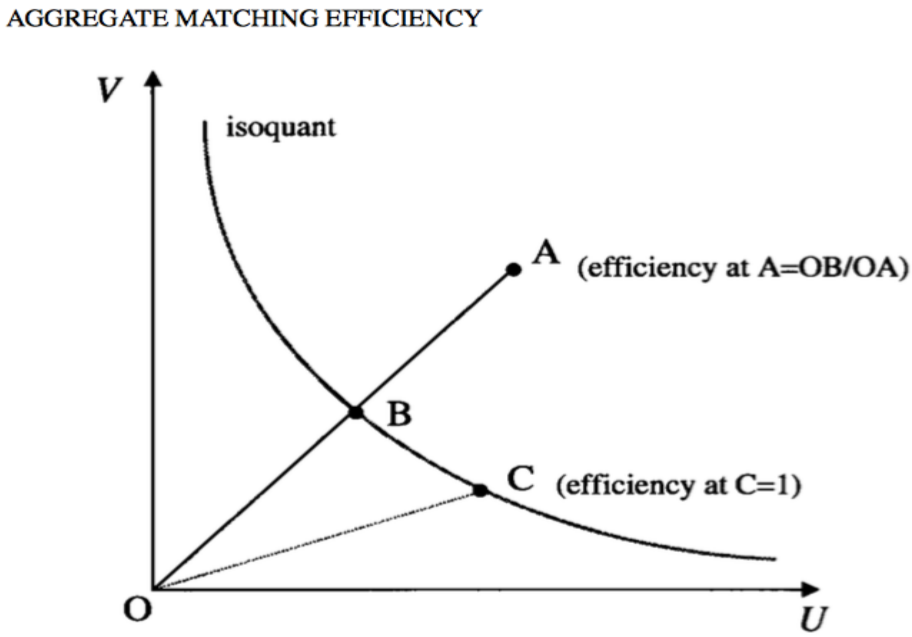

Figure 1: The efficiency frontier in the V-U space (Ibourk et al., 2004).

The stochastic frontier approach can be explained with help of Figure 1, where the downward-sloping curve represents an isoquant. The isoquant shows for every value of input factor U, the minimum amount of input factor V that is needed to produce the desired level of output factor M. For a given output level, point B and C correspond to two efficient input combinations, whereas point A represents an inefficient input combination. The so-called technical inefficiency for a producer operating at point A is defined as 1 minus the ratio OB/OA. According to the efficiency frontier, all observed input combinations, at a given level of output, must be either on or to the right of the isoquant seen in Figure 1. This implies that technical efficiency can at most be equal to 1. The stochastic frontier approach can be used to study how far the observed matching outcome is from the efficiency frontier and to test which factors that contribute to potential inefficiencies (Ibourk et al., 2004).

3 Literature Review

3.1 The Matching Process in the Swedish Context

From an international perspective, the matching function has been extensively studied in the empirical literature (see Petrongolo and Pissarides, 2001). The data usually varies from regional panel data to aggregated time series, but occupational data has also been studied. The empirical evidence on the matching function and matching efficiency in the Swedish labour market is, on the other hand, relatively scarce. Some of the most comprehensive and related studies are presented below. All studies are primarily based on data from the Swedish Public Employment Service.

The paper by Aranki and Löf (2008) is closest related to our thesis when it comes to the research topic. Their paper also analyses the regional matching process in Swedish labour market using a stock-flow matching function but for a different time period, i.e. 1992-2007. Aranki and Löf (2008) find that all input variables have a significant positive effect on the number of matches, but that the stock of unemployed has a larger impact than the stock of vacancies. Furthermore, the inflow of vacancies is found to be relatively more important than respectively stock. Their findings also indicate that matching efficiency varies among regions and, on average, is lower in the most densely populated regions compared to the less densely populated. The most important drawback of the paper by Aranki and Löf (2008) is that the empirical strategy does not allow matching efficiency to vary over time. We solve this problem by using the stochastic frontier approach. Moreover, our thesis gives a more comprehensive picture of the regional matching process as we extend the model by including variables that capture worker heterogeneity.

Two additional papers that give a good overview of the functioning of the Swedish labour market during more recent years are Marthin (2012) and Håkansson (2014). Similar to our thesis, both papers base their analysis on matching function theory but they use other empirical methods. Marthin (2012) focuses on measuring Swedish labour market imbalances using mismatching indices. According to Marthin (2012), geographical and occupational mismatch on the Swedish labour market can explain a significant part of total unemployment. The result also indicates that the occupational mismatch may be larger than the geographical mismatch. Furthermore, the occupational mismatch has changed size and composition over the years and might be an increasing problem. The study also shows that the occupational mismatch index is less for the inexperienced job seekers than for the experienced.

The research focus of Håkansson (2014) also differs from this thesis in the sense that the author studies the Swedish labour market from a national perspective and estimates job finding rates. Håkansson (2014) argues that the fact that the number of vacant jobs has increased during the last years, while job finding opportunities for unemployed have lagged behind, could be a sign of deteriorated labour market efficiency. Similar to our thesis, Håkansson (2014) is interested in the composition of the unemployment group. She states that deteriorated efficiency could be partly explained by changes in the composition of the unemployment group, where a larger share of unemployed belongs to vulnerable labour market groups. The change in composition is probably explained both by purely demographic developments, but also by increased labour participation among individuals with a weak attachment to the labour market during recent years. Furthermore, Håkansson (2014) finds that the job finding rate for the non-vulnerable group has recovered after the crisis, but that the job finding rate for vulnerable groups remains low. Therefore, an increase in the share of the vulnerable group among the unemployment stock has led to slow recovery of the unemployment rate.

A potential drawback with the studies of Marthin (2012) and Håkansson (2014) is that their results are derived from matching functions solely based on random matching. Following the findings of Forslund and Johansson (2007), who conclude that the Swedish labour market is characterized by stock-flow matching rather than random matching, this could be a rather simplified assumption.

3.2 International Studies Based on Stochastic Frontiers

Internationally, there exist a couple of papers among the stochastic frontier literature, which all are related to our thesis as we use similar empirical approaches to study regional labour markets. Both our thesis and the following papers, use the same definition of the inefficiency term in the stochastic model. This definition is based on the framework of Battese and Coelli (1995) and it makes it possible to analyse the determinants of labour market inefficiency. Our thesis differs from the following papers, as we are the first to use this empirical method to study Sweden. We also cover a more recent time period, in which the latest global financial crisis is included. Furthermore, in contrast to the other papers, our thesis is not only focused on random matching, but also stock-flow matching when we study the matching process.

Ilmakunnas and Pesola (2003), Ibourk et al. (2004), and Fahr and Sunde (2002, 2006) use the stochastic frontier approach to study regional labour markets in European countries during the 80's and 90's. They all find that there exist cross-regional variation in efficiency and that the elasticity of matches with respect to the unemployment stock is larger than the elasticity of the vacancy stock. Otherwise, the papers somewhat differ in research focus and results. Ilmakunnas and Pesola (2003) study the Finnish labour market and find evidence of constant returns to scale in the matching process. Young as well as long-term unemployed job seekers are, somewhat surprising, found to positively influence efficiency. Moreover, unemployment spillovers from neighbouring areas are found to have a negative effect on the matching efficiency. They also find pro-cyclical variation and a negative trend in the matching efficiency, indicating that business cycles have a large impact on regional efficiency estimates.

Ibourk et al. (2004) study French regions and show that aggregate matching efficiency has decreased during their research period. By disaggregating the pool of job seekers, the authors find that the share of the young individuals, females and immigrants increase efficiency in the matching process. The authors argue that these groups might be more prone to take temporary jobs, thereby having a positive impact on the matching. Furthermore, population density is found to have a positive effect on matching efficiency, whereas long-term unemployment has a negative effect.

The paper of Fahr and Sunde (2002, 2006) is based on German labour markets. Their results indicated that the efficiency ranking of regions is stable over time, but that efficiency has increased over time as a result of the German reunification. Age and the educational structure of the labour force are found as important determinants of matching efficiency. Moreover, the probability of matching is positively affected by the search intensity and competition among job seekers and firms.

Finally, our thesis is most similar to Hynninen (2009) and Hynninen et al. (2009) as we use the same choice of stochastic frontier model. Like Ilmakunnas and Pesola (2003), they study regional labour markets in Finland but for the years 1995-2004. In contrast to the papers of Ilmakunnas and Pesola (2003), Ibourk et al. (2004), and Fahr and Sunde (2002, 2006), they use an extended version of the stochastic frontier model. This version has the advantage of being able to separate inefficiency in the matching process from cross-sectional time-invariant heterogeneity (Greene, 2005a,b). In accordance with our thesis, Hynninen (2009) focuses on the composition of the job seeker stock. The author finds that the composition is an important explanation of matching efficiency. Especially the share of highly educated job seekers and the share of job seekers out of the labour force are found to increase efficiency. Long-term unemployed, on the other hand, is found to decrease efficiency. Furthermore, Hynninen et al. (2009) argue that regional inefficiencies are an important factor of total unemployment. They find that if the LLOs could produce their optimal amount of matches, unemployment would be 2.4% lower. Furthermore, if all LLOs had the same structure of job seekers as the most efficient one, unemployment would decrease with additional 1.4%. This indicates that inefficiencies in practices and management processes in regional LLOs may be more important than variations in structural factors.

4 Data

The data used in the empirical analysis is derived from the Swedish Public Employment Service. It comprises monthly labour market data for the Swedish counties for the time period 2006:1-2016:12. The chosen time period enables us to study potential changes in the matching efficiency following the financial crisis in 2008. Using regional data makes it possible to take into account potential regional differences in the labour market. The data contains both stocks and flows of registered unemployed job seekers as well as vacancies reported to the LLOs.

Empirical studies on the matching function differ in terms of how the number of matches in the labour market is measured. Since this strongly affects the implications of the empirical analysis, the choice of the dependent variable is very essential. Commonly used proxies for the number of matches are the total outflow from unemployment, the outflow from unemployment to employment and the flow of filled vacancies (Hynninen et al., 2009). In this study we choose to focus on the outflow from registered unemployment to employment as the dependent variable. Unlike the total outflow from unemployment, this measure does not include individuals who have left unemployment because of labour force withdrawal. The outflow from unemployment to employment is therefore less likely to vary in size over business cycles and across regions (Burgess et. al, 2001). Moreover, since the purpose of this study is to study the matching process from the supply side of the labour market, the flow of filled vacancies as the dependent variable is not relevant for this study.

According to the definition used by SPES, the stock of unemployed is referred to as the number of registered unemployed job seekers, including individuals who are actively searching for employment as well as individuals participating in labour market programmes. The outflow from unemployment to employment refers to the number of registered unemployed job seekers who are newly hired and has therefore left the stock of unemployment by the end of the month. The inflow to unemployment is measured as the number of persons who are newly registered as unemployed job seeker during the month. Since it is mandatory to register at the SPES to qualify for unemployment benefits, there is a strong incentive for unemployed job seekers to register. The SPES statistics also include the stock of vacancies, including the total vacancies at the end of the previous month and the inflow of new vacancies during the month.

Nevertheless, it is necessary to emphasize that the registered unemployed job seekers does not cover the total population of unemployed job seekers. According to Gottfries and Stadin (2016) the number of job seekers registered at the Public Employment Service has declined over time and was only 54% of the unemployment population in the Labour Force Survey by 2011. Another problem that arises is that far from all vacancies are reported to the SPES. The reason for this could be that employers are no longer required to report their vacancies to the SPES and may find it more efficient to use other recruiting channels. Aranki and Löf (2008) find that SPES had a market share of only 30-45 per cent in the 1990s and 2000s. This figure tends to be higher in times of great demand for labour than during recessions.

Against this background, the SPES statistics on unemployment and vacancies may be imperfect measures. However, these problems are not exceptional to this study but rather common in the empirical literature on labour market matching. Even though the statistics from SPES do not represent the total population of unemployed and vacancies in the economy, the statistics can be considered as a reliable indicator to determine the condition of the Swedish labour market (see Gottfries and Stadin, 2016; Håkansson, 2014; Marthin, 2012). Moreover, by comparing SPES data from 1992-2011 to data from the quarterly Labour Force Survey conducted by Statistics Sweden, Gottfries and Stadin (2016) find that both registered unemployment and vacancies, on aggregated level, fluctuated in the same way.

Besides the stock and flow variables in the stochastic frontier matching function, our empirical analysis includes control variables that capture labour market heterogeneity and that are considered to influence the matching efficiency (Fahr et al., 2006; Ibourk et al., 2004; Ilmakunnas and Pesola, 2003; Petrongolo and Pissarides, 2001). The inefficiency term in our stochastic frontier model is described as a function of variables describing the composition of the unemployment stock according to age, labour market status and education. Hynninen et al. (2009) argue that the age structure of the unemployment stock may have an effect on the matching efficiency due to different search intensity and reservation wages depending on age. Moreover, employers may treat the job seekers differently based on age, as age is commonly used as a proxy for experience. Thus, the proportion of unemployed job seekers below 25 and over 55 years of age are included in the matching function.

Unemployment duration is commonly used to control for ranking and variations in search intensity among job seekers. The share of long-term unemployed job seekers (over 1 year) is expected to have a negative effect on matching efficiency since prolonged unemployment tend to decrease the search intensity in addition to loss of skills (Barnichon and Figura, 2011). The share of unemployed in active labour market programmes (ALMP) is included as a proxy for regional policies determining the efficiency of the matching process. According to Ibourk et al. (2004), participants in ALMPs tend to have a harder time finding jobs on the open labour market and is therefore ALMP also expected to have a negative effect on matching efficiency.

We also study the effect of educational mismatch by using data on unemployment by educational attainment level. *Primary* refers to nine years of compulsory education, *secondary* refers to a completed upper secondary school, and *high* refers to post-secondary education of three years or more. The proportions of primary educated as well as highly educated unemployed job seekers are included in the model. This is motivated by the assumption that prolonged unemployment is less likely among higher educated job seeker. This means that highly educated job seekers are expected to have a positive impact on the matching efficiency.

Finally, the share of foreign-born individuals in the unemployment stock is included for each region. Since the 1990s, the labour migration from European countries to Sweden has shifted to primarily become asylum seekers from non-EEA countries. Because of poor labour market integration of immigrants, this is another disadvantaged group in the Swedish labour market. It is therefore highly relevant to control for the effect of this group on the matching efficiency in Sweden (Zakova, 2013).

Table 1. Descriptive statistics by region

Region	M	U	V	LTU	ALMP	Share <25	Share >55	Primary	High	Migrant	Hiring rate	Tightness
Stockholm	4 695	60 741	13 129	0.44	0.32	0.13	0.16	0.28	0.35	0.51	0.08	0.22
Uppsala	760	9 166	1 269	0.41	0.37	0.21	0.16	0.26	0.30	0.33	0.08	0.14
Södermanland	777	12 621	1 159	0.56	0.43	0.22	0.16	0.34	0.19	0.37	0.06	0.09
Östergötland	1 207	18 347	1 930	0.56	0.46	0.23	0.16	0.30	0.20	0.32	0.07	0.11
Jönköping	825	10 411	1 861	0.45	0.41	0.24	0.16	0.33	0.19	0.39	0.08	0.18
Kronoberg	465	6 756	1 104	0.49	0.42	0.24	0.14	0.32	0.23	0.42	0.07	0.16
Kalmar	711	8 208	1 604	0.42	0.45	0.25	0.18	0.31	0.19	0.27	0.09	0.20
Gotland	182	2 156	309	0.43	0.43	0.24	0.19	0.26	0.20	0.11	0.08	0.14
Blekinge	504	7 084	694	0.49	0.47	0.26	0.16	0.29	0.20	0.27	0.07	0.10
Skåne	3 589	51 422	5 602	0.47	0.40	0.20	0.15	0.28	0.26	0.40	0.07	0.11
Halland	770	8 983	1 133	0.43	0.39	0.22	0.18	0.28	0.23	0.31	0.09	0.13
Västra Götaland	4 299	56 712	7 806	0.49	0.40	0.21	0.15	0.28	0.25	0.38	0.08	0.14
Värmland	1 044	10 896	1 199	0.40	0.47	0.25	0.15	0.26	0.19	0.22	0.10	0.11
Örebro	914	11 796	1 133	0.49	0.43	0.24	0.15	0.33	0.18	0.35	0.08	0.10
Västmanland	766	10 972	1 114	0.50	0.43	0.24	0.14	0.31	0.19	0.36	0.07	0.10
Dalarna	960	9 374	1 259	0.39	0.41	0.23	0.17	0.31	0.18	0.24	0.10	0.13
Gävleborg	1 090	14 349	1 270	0.51	0.49	0.25	0.15	0.33	0.16	0.26	0.08	0.09
Västernorrland	882	11 187	1 050	0.45	0.44	0.24	0.17	0.28	0.19	0.23	0.08	0.09
Jämtland	526	5 135	565	0.37	0.47	0.25	0.16	0.26	0.20	0.20	0.10	0.11
Västerbotten	902	9 700	1 092	0.40	0.44	0.28	0.14	0.23	0.24	0.21	0.09	0.11
Norrbotten	1 042	10 689	1 228	0.44	0.47	0.25	0.17	0.25	0.18	0.20	0.10	0.11
Average	1 281	16 510	2 262	0.46	0.43	0.23	0.16	0.29	0.21	0.30	0.08	0.14

Notes: M denotes the number of matches during a month. U and V denote the number of unemployed job seekers and vacancies at the beginning of a month, respectively. Share LTU denotes the share of long-term unemployed in the unemployment stock, share ALMP denotes the share of unemployed participating in labour market programmes. Share <25 years and share >55 years denote the shares of the age groups in the unemployment stock. Primary and High denotes the unemployed by level of educational attainment, as a share of the unemployment stock. Migrant denotes the share of foreign-born individuals in the unemployment stock. Hiring rate is calculated by H/U and tightness by V/U.

Table 1 summarises the labour market statistics across regions. All figures are averages from 2006 to 2016. During an average month there are 16 510 unemployed job seekers and 2 262 vacancies registered at the regional SPESs. The outflow from unemployment to employment is on average 1 281, being highest in the most densely populated regions. Stockholm is the largest county in Sweden based on population and has an unemployment stock that is almost four times higher than the average across regions. Gotland, on the other hand, has an unemployment stock corresponding only to one-eighth of the average stock.

Table 1 does also display the structural variables that are used to assess the technical efficiency across regions. Although the variables vary across regions, most of the values vary within a small range. A large share of the unemployment stock is either long-term unemployed job seekers or participants in labour market programmes. This means that many of those registered at the SPES as unemployed may be individuals having a hard time finding jobs on the open labour market. Migrant job seekers account, on average, for 30% of the unemployed job-seeker stock. In educational composition, primary educated job seekers account for 29% on average, while highly educated job seekers account for 21%. This means that the largest group based on educational level have attained at least upper secondary education. During the research period, 23% of the unemployed are below 25 years and 16% are over 55 years.

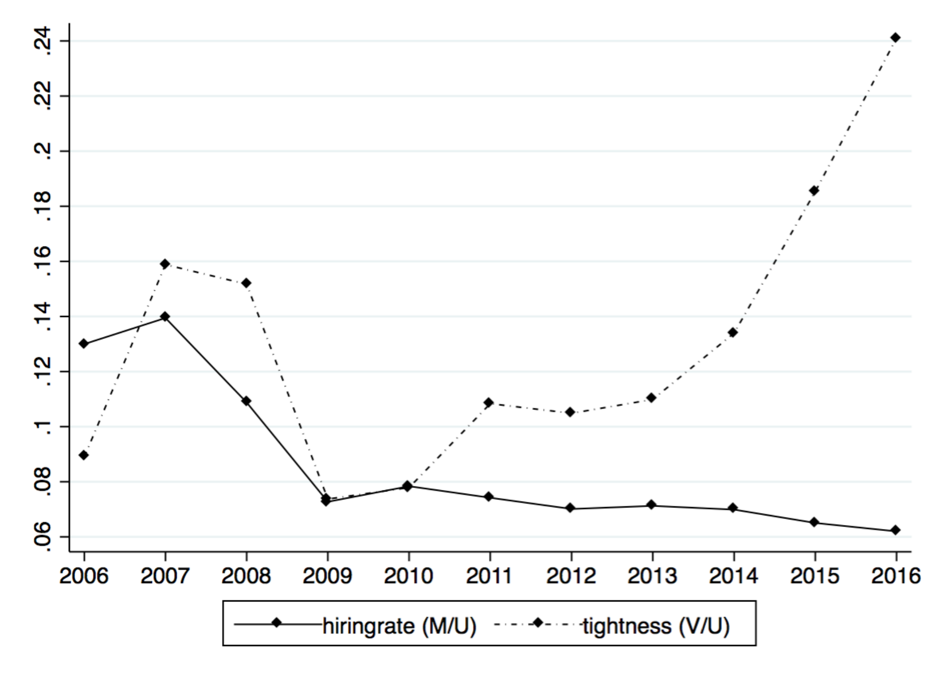

Figure 2: Hiring rate and labour market tightness over years.

Hiring rate measures the relationship between outflow of unemployment and unemployment stock and is calculated by the number of matches divided by the unemployment stock. The hiring rate in Swedish labour market is only 8% on average during the research period. Although this measure gives a glimpse of regions' ability to create labour market matches, one should bear in mind that this figure only includes the outflow from unemployment to employment registered at SPES. The labour market tightness is calculated by dividing the vacancy stock with the unemployment stock. Labour market tightness is on average 14%, but it varies from 9% to 22% between regions. Although the hiring rate and labour market tightness varies across regions, the variables are normally distributed, as the averages are not far away from each other. As seen in Figure 2, the hiring rate and labour market tightness showed an upward trend during 2006. After reaching a peak in 2007, both factors sharply declined by 50% during the recession. The years following the recession, the labour market tightness increased remarkably, while the hiring rate kept declining at a slow rate. As a result, the gap between the hiring rate and the labour market tightness has widened over the research period, which implies deterioration in the matching process. This serves as a starting point for utilizing a stochastic frontier analysis to estimate the changes in matching efficiency at the Swedish labour market.

5 Method and Empirical Model

When specifying our model, we assume that the matching function can be described by a Cobb-Douglas production function. This is in line with most of the empirical literature on labour market matching (see Petrongolo and Pissarides, 2001). We use logarithmic transformation in order to transform the Cobb-Douglas form into a log-linear function. Our stochastic frontier model can then be described in the following way:

$$\ln M_{i,t} = (\mu_i + \alpha \ln U_{i,t-1} + \beta \ln V_{i,t-1} + year + month) + v_{i,t} - u_{i,t} \tag{5}$$

where $M_{i,t}$ represents the number of matches (outflow from unemployment to employment) during a month t in county i, $U_{i,t-1}$ is the stock of unemployed, and $V_{i,t-1}$ is the stock of vacancies at the beginning of each month. We use the lagged values of U and V since the data of these variables are measured in the end of the month. $year$ and $month$ represent time dummies. To avoid perfect multicollinearity i.e. the dummy variable trap, we omit one year and one month dummy. The model described by equation (5) represents a matching function that only depends on the stocks of unemployed and vacancies, meaning that we assume random matching. Since we also have data of the inflows of unemployed and vacancies during each month, we are able to test for both random matching and stock-flow matching on the Swedish labour market. The model in (5) can then be extended so that the matching process instead is explained by stock-flow matching. The only difference is that the inflows of unemployed $\dot{U}_{i,t}$, and vacancies $\dot{V}_{i,t}$, are added to the model. The stochastic production frontier model can be rewritten as:

$$\ln M_{i,t} = (\mu_i + \alpha_1 \ln U_{i,t-1} + \beta_1 \ln V_{i,t-1} + \alpha_2 \dot{U}_{i,t} + \beta_2 \dot{V}_{i,t} + year + month) + v_{i,t} - u_{i,t} \tag{6}$$

According to Hynninen (2009) the expression in parentheses can be defined as the matching frontier, which gives the maximum output (i.e. number of matches) by the given quantities of the input factors (i.e. the number of unemployed and vacancies). μ_i denotes the time-invariant heterogeneity in the production of matches across counties.

In a stochastic frontier model, the error term consists of two parts so that $\varepsilon_{i,t} = v_{i,t} - u_{i,t}$. $v_{i,t}$ can be seen as a normal error term, which is usually assumed to be iid and follow the $N(0, \sigma_v^2)$ distribution. $u_{i,t}$ is a non-negative random variable that captures the technical inefficiency in the production process. Furthermore, $u_{i,t}$ is assumed to be distributed independently of $v_{i,t}$ and follows the $N(Z_{j,it}\delta_j, \sigma_u^2)$ distribution truncated at zero (Hynninen, 2009).

In accordance with the definition of Battese and Coelli (1995), the inefficiency term $u_{i,t}$ is a function of some inefficiency regressors and can be written as $u_{i,t} = Z_{j,it}\delta_j + w_{i,t}$. $Z_{j,it}$ is defined as the vector of inefficiency regressors, j stands for the number of inefficiency regressors, and δ_j represent the coefficients that will be estimated. The term $Z_{j,it}$ captures environmental factors that are believed to affect the efficiency. These factors can vary between cross-sectional units and time. In our case, we use variables describing the composition of the unemployed workforce as determinants of the regional inefficiency. These variables are described in section 4.

The last part of the inefficiency term $w_{i,t}$ is a random variable following a truncated normal distribution with zero mean, variance σ_u^2 and truncation point at $-Z_{j,it}\delta_j$. This follows from the fact that $u_{i,t}$ is a non-negative random variable. The result of these assumptions is that the variables describing the composition of the unemployed workforce influence the degree of technical efficiency, but not the shape of the production technology. This stands in contrast to the conventional fixed effects model (Battese and Coelli, 1995; Coelli et al., 1999). On this basis, we define the inefficiency function that we want to estimate together with model (5) and (6) as:

$$u_{i,t} = \delta_0 + \delta_1\left(ShareLTU_{i,t-1}\right) + \delta_2\left(ShareALMP_{i,t-1}\right) + \delta_3\left(Share < 25_{i,t-1}\right) +$$
$$\delta_4\left(Share > 55_{i,t-1}\right) + \delta_5\left(SharePRIMARY_{i,t-1}\right) + \delta_6\left(ShareHIGH_{i,t-1}\right) +$$
$$\left(ShareMIGRANT_{i,t-1}\right) + w_{i,t} \quad (7)$$

Furthermore, the inefficiency term and the parameters of the stochastic frontier are jointly estimated using the maximum likelihood estimator. The conditional estimates of the technical efficiency can then be obtained by:

$$TE_{i,t} = [\, exp\,(-u_{i,t}) |\, M, U, V, Z\,] \quad (8)$$

This measure gives the absolute technical efficiency. Hence, it is not a relative measure to the best or worst in the sample. This means that if the technical efficiency equals 1, the number of matches is on the frontier. In all other cases, the technical efficiency must be less than 1 (Hynninen, 2009).

The model described above is usually called the "true" fixed-effects model in the stochastic frontier literature. This is since the time-invariant heterogeneity μ_i is separated from the inefficiency term $u_{i,t}$. A separation between the two terms is not possible in conventional fixed-effects models (Greene, 2005a, b). We estimate the "true" fixed-effects stochastic frontier model via the Stata command *sfpanel*, which estimates stochastic frontier models for panel data using the maximum likelihood estimator. The command also allows us to address potential heteroscedasticity in the inefficiency component, by defining the variance as a function of the inefficiency regressors (Belotti et al., 2012). For the sake of comparison, we also estimate our models using the conventional panel data methods random-effects and fixed-effects. In this case, inefficiency regressors are included in the models as ordinary control variables and not as a function of the inefficiency term.

6 Results

6.1 Matching Analysis

Table 2. Matching process estimates based on non-normalized data

Variables	Conventional panel data			Stochastic frontier	
Dependent variable: ln M_t	Random (1)	Fixed (2)	Fixed (3)	True fixed (4)	True fixed (5)
ln U_{t-1}	0.891***	0.691***	0.712***	0.708***	0.718***
	(0.025)	(0.052)	(0.056)	(0.052)	(0.054)
ln V_{t-1}	0.087***	0.058*	0.028	0.035**	0.015
	(0.024)	(0.032)	(0.022)	(0.017)	(0.016)
ln $\dot{U}_t$			-0.149*		-0.075*
			(0.074)		(0.045)
ln $\dot{V}_t$			0.072***		0.083***
			(0.017)		(0.017)
Inefficiency controls					
(Share LTU)$_{t-1}$	-0.447***	-0.217*	-0.223*	0.347**	0.364**
	(0.097)	(0.114)	(0.114)	(0.176)	(0.185)
(Share ALMP)$_{t-1}$	0.150	0.153	0.114	-0.184	-0.135
	(0.184)	(0.157)	(0.151)	(0.164)	(0.158)
(Share <25)$_{t-1}$	-0.437	-0.597	-0.573	0.872***	0.751**
	(0.447)	(0.447)	(0.363)	(0.337)	(0.331)
(Share >55)$_{t-1}$	1.002	-0.071	-0.264	0.109	0.275
	(0.634)	(0.610)	(0.604)	(0.708)	(0.647)
(Share PRIMARY)$_{t-1}$	1.227***	0.791**	0.475	-1.035***	-0.856***
	(0.319)	(0.306)	(0.276)	(0.330)	(0.318)
(Share HIGH)$_{t-1}$	1.253***	0.406	-0.021	-0.899***	-0.517
	(0.307)	(0.326)	(0.361)	(0.341)	(0.387)
(Share migrant)$_{t-1}$	-1.225***	-1.278***	-1.156***	1.588***	1.525***
	(0.291)	(0.302)	(0.322)	(0.221)	(0.247)
Constant	-2.118***	0.401	1.130***	0.739***	0.592***
	(0.471)	(0.826)	(0.395)	(0.231)	(0.225)
Returns to scale	0.978	0.749***	0.664***	0.743***	0.722***
Observations	2,772	2,772	2,772	2,772	2,772
Number of regions	21	21	21	21	21
R^2		0.830	0.837		
Hausman, Chi^2		424.65***			
F-test			59.25***		
LR-test					94.22***
Log-likelihood				1887.83	1934.94

Notes: The models include yearly and monthly dummies in the function. Clustered standard errors are reported in parentheses. *** denote statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. In tests for returns to scale, *** denote deviation from unity at the 1% level. Hausman test favours the fixed-effects model at the 1% level.The F-test rejects the hypothesis that model 3 is nested in model 2 at the 1% level. The LR-test rejects the hypothesis that model 5 is nested in model 4 at the 1% level.

The main regression results are presented in Table 2. The first three columns are estimated with traditional random-effects and fixed-effects respectively. The last two columns represent the stochastic frontier analysis, which is estimated with Greene's (2005a, b) "true" fixed-effects. All specifications are regressed with monthly and yearly dummies in order to control for seasonal and yearly effects. We define our base group as unemployed job seekers aged between 25 and 55 years with an unemployment spell shorter than a year. With respect to education, secondary educated job seekers form the base group. The estimates are studied in relation to the base group. All specifications, except from the one estimated with random-effects, exhibit decreasing returns to scale. Yet, in the model with random effects we cannot reject the null hypothesis, indicating that the model exhibit constant returns to scale.

Column 1 and 2 show the results for a random matching process, which only depends on the stocks of unemployed and vacancies. Column 1 is estimated with random-effects, whereas Column 2 is estimated with fixed-effects. When it comes to the two main input factors, U and V, we can see that unemployed job seekers are relatively more important than vacancies for the number of matches. More precisely, the result of the random-effects estimator indicates that the number of matches increases with 0.89% if the stock of unemployed increases with 1% and with 0.09% if the stock of vacancies increases with 1%. The result from the fixed-effects estimator shows lower elasticities, namely the number of matches increase with 0.69% if the stock of unemployed job seekers increases with 1%. The effect of a one-percentage point increase in the stock of vacancies is equal to a 0.06% increase in the number of matches. The difference in importance of the two main input factors is consistent with previous empirical literature, both in the international and the Swedish context (e.g. Aranki and Löf, 2008; Hynninen et al., 2009; Petrongolo and Pissarides, 2001 etc.).

Regarding our inefficiency control variables, we find that the coefficients for $ShareLTU$, $SharePRIMARY$, $ShareHIGH$ and $ShareMIGRANT$ are statistically significant at the 1% level in the model estimated with random-effects. In the model estimated with fixed effects, the coefficient for $ShareHIGH$ is statistically insignificant at the 10% level. For the three other statistically significant coefficients, the signs are the same as in the random-effects model. Namely, $SharePRIMARY$ and $ShareHIGH$ have a positive impact on the number of matches, whereas $ShareLTU$ and $ShareMIGRANT$ have a negative impact on the number of matches.

We perform a Hausman test to conclude which of the two models that is most appropriate. The Hausman test favours the fixed-effects model at the 1% level. This is in line with what we expected, as it is reasonable to believe that the regional-specific effect is correlated with the explanatory variables in this case. Hence, the assumption for random effects could not be fulfilled. Otherwise, the random effects model would be used, as it is more efficient than fixed effects in that case (Stock and Watson, 2011).

Column 3 presents the result from the model with stock-flow matching. Based on the Hausman test statistics, we decide to only report the result for the fixed-effects model. The model with stock-flow matching indicates that the inflow of vacancies has a greater impact on the number of matches than the stock of vacancies, as the coefficient for the vacancy stock is statistically insignificant at the 10% level. Furthermore, the inflow of unemployed has a small but negative effect, which makes interpretation difficult since it goes against the stock-flow theory and the findings of Aranki and Löf (2008). On the other hand, our data is based on different time period, which covers some years of economic recession. Gottfries and Stadin (2016) also find negative, but statistically insignificant, coefficients for the unemployment inflow when using SPES data and studying a time period that overlaps ours. Since the coefficient only is statistically significant at the 10% level, we should probably interpret it with caution in this case. Aranki and Löf (2008) argue that a relatively common result in the literature of stock-flow matching is to find that the flow variables positively influence the number of matches. Their results of the Swedish matching process suggest that both stocks and inflows have a statistically significant positive effect on the number of matches, but that the stock of unemployed and the inflow of vacancies have the largest impact. The result of Coles and Smith (1998), on the other hand, is more in line with our results. Namely, that only the inflow of new vacancies, and not the stock of vacancies, increases the probability of unemployed to find a job. $ShareMIGRANT$ is the only structural variable that is statistically significant at conventional significance levels in Column 3.

For a comparison of the model in Column 2 and 3, we perform a F-test with the null hypothesis that the stock-flow matching model is nested in the model with random matching. The test result rejects the null hypothesis, indicating that the flow variables are of importance.

Column 4 presents the stochastic frontier analysis estimated with a random matching process. When it comes to stock elasticities, the model indicates the same tendencies as in previous results. The model with random matching gives that the elasticity of matches with respect to unemployment is larger than the one for vacancies.

In the stochastic frontier framework, the sign of the remaining explanatory variables should be interpreted as having a negative or a positive effect on the inefficiency in the production process of matches. From Column 4, we find that long-term unemployment increases matching inefficiency. A one-percentage point increase in the long-term unemployment group increases inefficiency by 0.35%. The result goes hand in hand with the ranking theory, stating that employers prefer short-term unemployed applicants to long-term unemployed applicants. The result is also in line with most of the empirical literature on labour market matching, where prolonged unemployment is expected to have a negative effect on matching efficiency due to decreased search intensity and loss of skills.

We also find that an increase in the share of young unemployed increases inefficiency in the matching process. This could for instance be a result of lower search intensity among youth, but also because young individuals are likely to accumulate less work experience. Furthermore, according to the results, an increase in the share of migrants in the unemployment stock has the largest negative impact on matching efficiency. A one-percentage point increase of this group leads to an increase in inefficiency of 1.6%. The negative sign is in line with our expectations as foreign-born unemployed are often considered to be a group with a vulnerable position at the Swedish labour market (Zakova, 2013). A negative impact on the number of matches could theoretically be both a result of lower search intensity or ranking due to language barriers, smaller labour market networks or discrimination.

The result in Column 4 also indicates that both the share of primary and highly educated unemployed have a negative impact on the inefficiency, hence larger shares of these groups increase the efficiency in the matching process. With some surprise we find that the coefficient for primary educated is higher than the coefficient for highly educated. However, these results are consistent with the findings of Fahr and Sunde (2006). They state that highly educated might have higher search intensity than other groups, resulting in higher matching efficiency. One could also argue that higher educated job seekers in the Swedish context may find it more efficient to use other recruiting channels since vacancies for temporary jobs and low skilled jobs are the most common type gathered by the SPES (Arbetsförmedlingen, 2011).

There are some potential explanations for the sign and coefficient of primary educated group. It could be that individuals with only primary education have lower reservation wages than other groups, thereby being more desirable for employers to hire compared to the base group of secondary educated individuals. Fahr and Sunde (2006) also mention that searching for applicants to low qualified work tasks might be less time-consuming, which leads to fewer frictions in these segments of the labour market. Another thing to have in mind when analysing educational data from the Swedish Public Employment Service is that the information is based on the unemployed individual's own statement (Statskontoret, 2015). Hence, educational information is not based on any registers or documentations and the results should therefore be evaluated with caution.

Column 5 shows the result for the stochastic frontier model, but this time defined by stock-flow matching. Compared to Column 4, we find that the coefficient for the vacancy stock now is statistically insignificant but that the coefficient for the unemployment stock is relatively robust. We also find that both inflow variables have a statistically significant effect on the matching process. The unemployment inflow has a small and negative effect on the number of matches, whereas the vacancy inflow has a small but positive effect. When it comes to all the other variables, we find that the model in Column 5 produce very similar results to the previous model. The only exception is that the coefficient for the group of highly educated now is statistically insignificant.

In order to compare the two models in Column 4 and 5, we perform a likelihood-ratio test. The test rejects the hypothesis that the stock-flow model is nested in the model with random matching. Hence, we cannot assume that the coefficients for the flow variables equal zero.

Table 3. Matching process estimates based on normalized data

Variables	Conventional panel data			Stochastic frontier
Dependent variable: ln M_t	Random (1)	Fixed (2)	Fixed (3)	True fixed (4)
ln U_{t-1}	0.632***	0.664***	0.677***	0.695***
	(0.041)	(0.052)	(0.051)	(0.054)
ln V_{t-1}	0.056*	0.055*	0.024	0.013
	(0.029)	(0.032)	(0.021)	(0.015)
ln $\dot{U}_t$			-0.161**	-0.085*
			(0.076)	(0.047)
ln $\dot{V}_t$			0.070***	0.081***
			(0.017)	(0.017)
Inefficiency controls				
(Share LTU)$_{t-1}$	-0.301***	-0.221*	-0.225*	0.362**
	(0.092)	(0.11)	(0.11)	(0.176)
(Share ALMP)$_{t-1}$	0.227	0.161	0.119	-0.141
	(0.153)	(0.160)	(0.154)	(0.160)
(Share <25)$_{t-1}$	-0.771	-0.548	-0.509	0.709**
	(0.473)	(0.436)	(0.365)	(0.340)
(Share >55)$_{t-1}$	-0.512	-0.183	-0.433	0.384
	(0.547)	(0.577)	(0.568)	(0.620)
(Share PRIMARY)$_{t-1}$	0.374	0.806***	0.488*	-0.847***
	(0.261)	(0.276)	(0.261)	(0.298)
(Share HIGH)$_{t-1}$	0.057	0.431	0.009	-0.505
	(0.271)	(0.316)	(0.339)	(0.379)
(Share migrant)$_{t-1}$	-1.137***	-1.278***	-1.153***	1.507***
	(0.367)	(0.309)	(0.324)	(0.250)
Constant	-2.709***	-2.877***	-3.328***	0.576***
	(0.234)	(0.272)	(0.314)	(0.218)
Returns to scale	0.687***	0.719**	0.611***	0.704***
Observations	2,772	2,772	2,772	2,772
Number of regions	21	21	21	21
R^2		0.837	0.845	
Hausman, Chi^2		93.79***		
Log-likelihood				1933.41

Notes: The models include yearly and monthly dummies in the function. Clustered standard errors are reported in parentheses. *** denote statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. In tests for returns to scale, *** denote deviation from unity at the 1% level. Hausman test favours the fixed-effects model at the 1% level

In the empirical matching literature based on regional labour markets, it is common to consider differences in population density or population size among regions. Münich et al. (1998) argue that variations in the labour force size may lead to spurious scale effects when using unscaled variables. In this thesis, we follow the example of Aranki and Löf (2008) and normalize all variables with the respective region's population size. In this way we can check the robustness of our results. The regression results based on population-normalized data are presented in Table 3. The results are on average very similar to the ones in Table 2, indicating that there is no serious problem with spurious scale effects in our case. In Column 1, educational variables are now statistically insignificant, but otherwise, the signs and significance levels are the same. Many coefficients are slightly lower in Table 3 compared to their counterparts in Table 2. This also results in overall lower estimates of the returns to scale when we control for population size.

The result from the stochastic frontier analysis based on normalized data is presented in Column 4. We only report the result from the stock-flow matching model, as the maximum likelihood estimator did not converged for the model with random matching. Furthermore, the likelihood-ratio test in Table 2 indicated that the flow variables should be included in the model. The result for the stochastic frontier model, based on normalized data, is very close to result obtained with non-normalized data. All coefficients have the same sign and significance level, but the coefficients are overall smaller.

6.2 Efficiency Analysis

To investigate regional disparities and developments in matching efficiency, we continue our analysis with further use of the stochastic frontier approach. According to our knowledge, there are no previous studies of regional labour market matching in Sweden based on a stochastic frontier analysis. This makes the results interesting, but with the downside that there are few studies in the Swedish context to compare with. The paper by Aranki and Löf (2008) is closest related among the Swedish literature. Since their study is based on conventional fixed effects, they measure the matching efficiency as the time-invariant regional-specific effect. Ibourk et al. (2004) argue that a problem with this approach is that regional efficiency is assumed to be time-invariant, which could be an unrealistic assumption. Furthermore, the traditional fixed-effects model only provides efficiency estimates that are measured in relation to the best or the worst in the sample.

Table 4. Matching efficiency measures by region

Region	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Average efficiency
Stockholm	0.334	0.329	0.347	0.356	0.317	0.312	0.294	0.280	0.273	0.264	0.260	0.306
Kronoberg	0.427	0.390	0.358	0.342	0.355	0.297	0.288	0.305	0.285	0.250	0.236	0.321
Skåne	0.394	0.386	0.375	0.369	0.330	0.317	0.317	0.298	0.284	0.273	0.264	0.328
Västra Götaland	0.380	0.365	0.357	0.364	0.356	0.331	0.326	0.307	0.302	0.293	0.276	0.332
Södermanland	0.433	0.393	0.370	0.356	0.363	0.323	0.315	0.321	0.296	0.267	0.251	0.335
Jönköping	0.425	0.397	0.379	0.386	0.370	0.315	0.291	0.304	0.290	0.272	0.267	0.336
Västmanland	0.444	0.394	0.395	0.371	0.352	0.326	0.304	0.302	0.296	0.277	0.268	0.339
Örebro	0.390	0.346	0.335	0.356	0.367	0.347	0.343	0.346	0.326	0.315	0.294	0.342
Östergötland	0.448	0.401	0.368	0.378	0.365	0.338	0.324	0.322	0.306	0.302	0.279	0.348
Halland	0.406	0.434	0.412	0.416	0.377	0.368	0.358	0.343	0.329	0.313	0.291	0.368
Uppsala	0.426	0.431	0.413	0.418	0.373	0.381	0.345	0.338	0.324	0.322	0.296	0.370
Blekinge	0.498	0.473	0.422	0.409	0.418	0.373	0.343	0.348	0.333	0.320	0.299	0.385
Gävleborg	0.492	0.481	0.461	0.435	0.401	0.393	0.373	0.361	0.332	0.319	0.291	0.394
Kalmar	0.508	0.465	0.445	0.464	0.415	0.389	0.372	0.357	0.327	0.326	0.298	0.397
Västernorrland	0.490	0.459	0.440	0.428	0.396	0.394	0.419	0.390	0.362	0.341	0.323	0.404
Västerbotten	0.484	0.467	0.462	0.432	0.429	0.381	0.380	0.385	0.368	0.352	0.348	0.408
Norrbotten	0.503	0.476	0.453	0.457	0.407	0.385	0.380	0.368	0.355	0.362	0.353	0.409
Värmland	0.518	0.473	0.443	0.458	0.422	0.400	0.395	0.369	0.361	0.344	0.331	0.410
Dalarna	0.490	0.464	0.480	0.490	0.426	0.443	0.418	0.394	0.360	0.333	0.310	0.419
Jämtland	0.507	0.494	0.477	0.471	0.411	0.450	0.436	0.409	0.367	0.357	0.342	0.429
Gotland	0.626	0.542	0.564	0.541	0.489	0.458	0.459	0.435	0.428	0.415	0.413	0.488
Sweden	0.458	0.431	0.417	0.414	0.388	0.368	0.356	0.347	0.329	0.315	0.300	0.375

We present our results of technical efficiencies in Table 4. All estimates of technical efficiency are calculated in accordance with the definition of TE described in Section 5. Furthermore, the estimates are based on the regression output from the stochastic-frontier model with stock-flow matching and population-normalized data, seen in Column 4 in Table 3. We have compared these estimates to alternative technical efficiency estimates based on non-normalized data and random matching instead, but we found no clear differences among the results. From Table 4, we can see that the average efficiency in the sample is 0.375. According to the stochastic production frontier theory, this indicates that Swedish counties, on average, only produce 37.5% of their potential matches. However, it is difficult to evaluate if this efficiency estimate should be seen as relatively high or low since there are no obvious counterparts in this case. Furthermore, technical efficiency estimates tend to be quite sensitive to the distributional assumption of the inefficiency term (Belotti et al., 2012), which could make absolute inference problematic. In our case, we are more interested in the efficiency relation between regions and over time. Hence, further interpretation of the technical efficiency estimates *per se* is of second interest.

Table 4 also shows the average regional efficiency by year and for the whole research period. All counties are arranged according to their average efficiency over time, from the county with the lowest average efficiency to the county with the highest average efficiency. According to our estimates, efficiency varies between Swedish regions. Stockholm has the lowest matching efficiency, followed by Kronoberg, Skåne and Västra Götaland. It can also be seen that Gotland is the county with the highest average efficiency. Thereafter follow Jämtland, Dalarna and Värmland. Compared to Aranki and Löf (2008), our result differs in the sense that we have two other counties with the lowest and highest efficiency estimates, but our result shows the same overall trend. Namely, more densely populated counties have, on average, lower efficiency scores than less densely populated counties. Since Aranki and Löf (2008) cover the time period 1992-2007, and we study the years 2006-2016, this pattern of regional matching efficiency seems to be relatively robust over time and empirical strategies. Aranki and Löf (2008) argue that a potential explanation for this could be that new jobs that emerge in metropolitan regions might require skills that there is a shortage of. Furthermore, they argue that the bottlenecks that arise in these regions as a consequence of matching problems, in turn can create problems related to wage formation and employment growth that affect the country as a whole. This hypothesis is also supported by the findings of Kano and

Ohta (2005), which find a negative relationship between matching efficiency and population density in Japanese regions.

Our results contradict the findings from some other countries. For France, Ibourk et al. (2004) find a concave relationship between population and the efficiency. They argue that social networks and information transformation may increase with population density, resulting in more efficient matching as long as congestion effects do not dominate it. In West Germany, Fahr and Sunde (2002; 2005) find that major urban areas have the highest efficiency when neighbouring effects is included, indicating that many of the major cities attract workers from surrounding areas. For our results, this could implicate that Swedish densely populated may suffer from congestion effects or that it would be wise to account for spatial effects to not get biased results. On other hand, Swedish regions are quite large and cover much of the neighbouring effects from municipalities close to the largest cities.

Another important aspect that can be seen in Table 4 is that all 21 counties have lower technical efficiency estimates in 2016 than in 2006. The table shows a clear trend in declining efficiency. However, it can also be seen that some regions have seen larger declines than others. For instance, Kronoberg had by margin a relatively more efficient matching process than Stockholm in 2006, but by 2016 Kronoberg had become the county with the lowest efficiency.

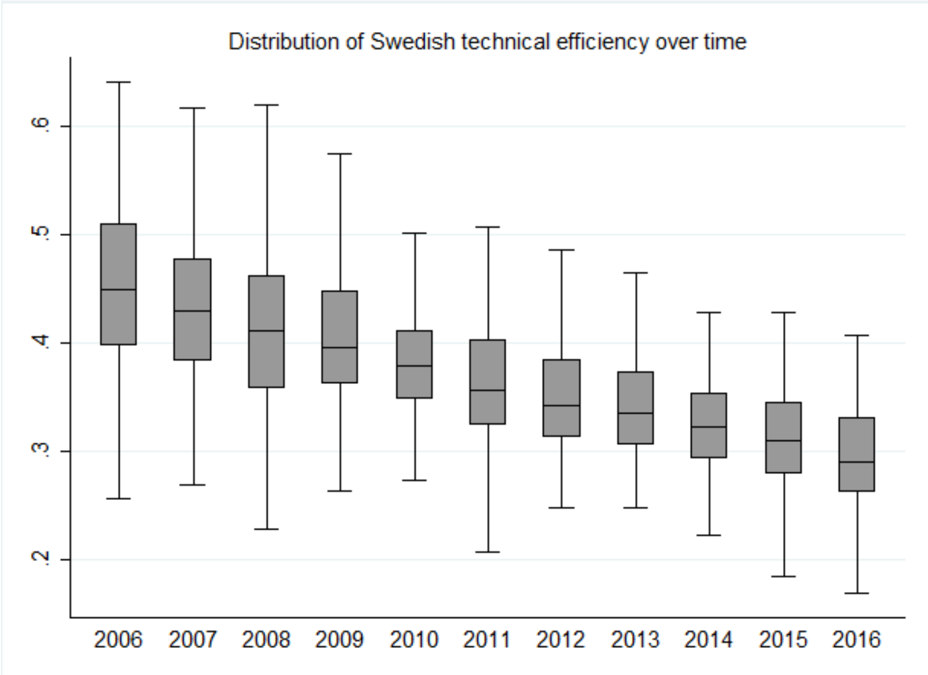

Figure 3: Box plot of the distribution of Swedish technical efficiency over time, where the vertical line within the box represent the median technical efficiency.[1]

Figure 3 shows the changes in matching efficiency in a more compact way. The negative trend is pronounced and the median efficiency seems to have decreased for each year that has passed. According to Figure 3, it also looks like the dispersion in efficiency among regions has declined over the research period. Our result is line with the view of the Riksbank, which during the last years has reported about deteriorated matching efficiency at the Swedish labour market (see for instance Håkansson, 2014 and Sveriges Riksbank, 2012). The result is also in line with much of the international literature in the field. For instance, OECD (2012) mentions that there are signs of deteriorated matching efficiency among a number of OECD countries after the crisis.

Håkansson (2014) argues that changes in the composition of the unemployment pool in some extent can explain increased inefficiency, as an increasing proportion of the unemployed belongs to groups with a weak attachment to the labour market. She mentions that the change of the composition partly can be explained by purely demographic developments, but also that new policy reforms as the earned income tax credit and changes of the unemployment and sickness insurance have been important. This based on the

[1] Based on the regional technical efficiency estimates in Table 4. The bottom and top of the boxes represent the first and third quartiles. The upper and lower whiskers represent the greatest and lowest values of the technical efficiency in the data set for each year when outliers are excluded.

assumption that the reforms probably have increased labour force participation among individuals who are closer to the periphery of the labour market.

Studying the changes of the composition of our unemployment stock over the research period, we find that most groups are relatively robust in size over time (see Table 5 in Appendix). However, there are distinct changes in the share of immigrants and long-term unemployed, which have grown in size. Among the share of immigrants, there is a quite steady growth during the whole period but among the share of long-term unemployed, we see a large increase during 2010-2011. Since the inefficiency regressors partly explain the technical efficiency, an increase in these groups could explain some part of deterioration in matching efficiency. If one believes that the financial crisis 2008 had a particular pronounced effect on these groups, the crisis may affect the deterioration in Swedish matching efficiency through these channels. The increase in long-term unemployed during 2010-2011 could be a natural consequence of several individuals losing their job during the crisis and struggling to find a new one. The steady increase in the share of immigrants could be explained by other reasons, such as increasing flows of refugees globally and Swedish immigration policies.

However, Håkansson (2014) also finds that the deterioration in labour market efficiency is less pronounced when data from the Labour Force Survey is used instead of data from the Swedish Public Employment Service. She argues that the two sources partly capture different kind of individuals and for this reason gives slightly different pictures.

The Riksbank (2012) shows evidence of weakened geographical mobility in Sweden, which until 2008 was strongly correlated with the annual change in the vacancy rate. Furthermore, Marthin (2012) argues that mismatching between occupations is an increasing problem, especially for experienced job seekers. Hence, increased occupational mismatch may be a potential explanation to deteriorated matching efficiency. Both decreased geographical mobility and increased occupational mismatches could be two important explanations for reduced efficiency on the labour market.

7 Conclusion

This paper investigates differences in regional matching efficiency as well as changes in efficiency that may have taken place following the financial crisis in 2008. The study is based on regional panel data from the Swedish Public Employment Service covering the time period 2006 to 2016. We utilize the ”true” fixed-effects stochastic frontier (see Greene 2005a, b) in order to control for time-invariant heterogeneity that may affect the technical inefficiency, which causes deviation from the frontier. The estimates have been compared with conventional fixed-effects estimates.

The main findings indicate large regional disparities in the Swedish matching efficiency. The results show that more densely populated counties, on average, have lower matching efficiency than less densely populated counties. A potential explanation could for instance be obstacles in geographical mobility. Regional factors may prevent job seekers from moving to a new location to obtain employment, such as regional disparities in the housing market. If regional employment patterns are driven by strong regional specific factors rather than by national factors, place-based policies targeting underperforming regions may be more adequate than national policies in order to dampen matching inefficiencies and to reduce labour market disparities across regions.

Furthermore, the regional matching efficiency in the Swedish labour market show a deteriorating trend during the last decade. It seems to be no visual structural break in the trend in conjunction with the financial crisis 2008. Since the deterioration started before the crisis, this could indicate that there are other more persistent structural imbalances at the Swedish labour market. Hence, it remains unclear if there is a causal relationship between deteriorating matching efficiency and the financial crisis. To determine whether there is an actual causal relationship is beyond the scope of this study and requires further investigation.

The empirical results show that inefficiencies in the Swedish labour market are determined by the composition of the unemployment stock. We find that the shares of the young and immigrants, as well as long-term unemployed job seekers have a negative impact on the matching efficiency. The shares of primary educated job seekers, on the other hand, has a positive impact. In order to better meet labour market and skill needs a policy priority should therefore be to further improve the integration of the young, foreign-born, and long-term unemployed job seekers. Studies have shown that vocational training and targeted measures towards job seekers with a weak attachment to the labour market significantly improve the likelihood to new matches and reduce the duration of unemployment (Forslund et al., 2011; SOU, 2011).

A drawback in this study is the structure of SPES's statistics. These statistics are based on individuals who are voluntarily registered as unemployed at SPES. This implies that the data may not be representative of the total population of unemployed job seekers and could therefore be a threat to the external validity of our results. Statistics from the Labour Force Survey (LFS), on the other hand, provide a more comprehensive and representative sample of the population of working age. However, LFS data at the regional level is unfortunately limited in coverage, leaving us with no choice but to use SPES statistics for the purpose of this study.

Another limitation with this thesis is that we ignore spatial effects in the matching process. It means that we do not take into account the potential influence of neighbouring counties' labour markets. Assuming that there are no interaction between regional markets are probably unrealistic, as labour markets neither are independent nor homogenous. Hence, omitting spatial spillovers may bias the results.

The potential limitations of our study result in some concrete suggestions of further research. For instance, spatial spillovers could be incorporated in the model by including neighbouring counties' unemployment and vacancy rate weighted by their respectively labour force. Another suggestion for further research is, to the extent it is possible, compare the regional matching function based on SPES data with a corresponding matching function based on LFS data from Statistics Sweden. Since LFS data is measured only on yearly basis, this would also open up for including other explanatory variables that may affect the inefficiency function or the matching function directly. Such a variables could for instance be measures of the composition of the vacancy stock, business cycles or geographic mobility.

Finally, we would also like to advocate for further use of the stochastic frontier approach when studying the Swedish labour market. This would enable comparisons between different stochastic frontier techniques and distributional assumptions, which could result in deepened understanding of labour market efficiency.

References

Andersson, P. M. & Burgess, S. M. (2000). Empirical Matching Functions: Estimation and Interpretation Using Disaggregate Data. Review of Economics and Statistics 82(1), 93-102.

Aranki, T. & Löf, M. (2008). The matching process on the Swedish labour market. Sveriges Riksbank Economic Review 2008:1.

Arbetsförmedlingen (2017-02-13). *Lägre arbetslöshet i de flesta län* [Electronic]. Available: https://www.arbetsformedlingen.se/Om-oss/Pressrum/Pressmeddelanden/Pressmeddelandeartiklar/Riket/2017-02-13-Lagre-arbetsloshet-i-de-flesta-lan.html [2017-05-21].

Arbetsförmedlingen. (2011). *Effective Services for Employers* [Electronic]. Available: http://ec.europa.eu/social/BlobServlet?docId=7556&langId=en.[2017-05-21]

Anxo, D. & Ericson, T. (2015). Labour Market Measures in Sweden 2008-13: The Crisis and Beyond. International Labour Office, Research Department, Geneva.

Barnichon, R. & Figura, A. (2011). What Drives Matching Efficiency? A Tale of Composition and Dispersion. Finance and Economics Discussion Series 2011:10.

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(2), 325–332.

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

Blanchard, O. & Diamond, P. (1994). Ranking, unemployment Duration and Wages. Review of Economic Studies 61(3), 417-434.

Boeri, T. & Burda, M. C. (1996). Central and Eastern European Labour Markets: Active Labour Market Policies, Job Matching and the Czech Miracle. European Economic Review 40, 805-17.

Coelli, T., Perelman, S. & Romano, E. (1999). Accounting for environmental influences in stochastic frontier models: with application to international airlines. Journal of Productivity Analysis 11(3), 251–273.

Coles, M.G. & Smith, E. (1998). Marketplaces and matching. International Economic Review 40(4), 851–868.

Dmitrijeva, J. & Hazans, M. (2007). A Stock–Flow Matching Approach to Evaluation of Public Training Programme in a High Unemployment Environment. Labour 21(3), 503–54.

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

Fahr, R. & Sunde, U. (2006). Regional Dependencies in Job Creation: An Efficiency Analysis for Western Germany. Applied Economics 3, 1193–1206.

Forslund, A., Fredriksson, P. & Vikström, J. (2011). What active labour market policy work in a recession? Nordic Economic Policy Review 2011 (1), 171–201.

Forslund, A. & Johansson, K. (2007). Random and stock-flow models of labour market matching – Swedish evidence. Working paper 2007:11, IFAU.

Gottfries, N. & Stadin, K. (2016). The Matching Process: Search Or Mismatch? Working paper 2016:4. Department of Economics, Uppsala University.

Greene, W. (2005a). Fixed and Random Effects in Stochastic Frontier. Journal of Productivity Analysis 23(1), 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. & Petrongolo, B. (2005). Stock-flow matching and the performance of the labor market European Economic Review 49(8), 1987-2011.

Hynninen, S-M. (2009). Matching Heterogeneous Job Seekers in Local Labor Markets: A Stochastic Frontier Approach. Journal of Productivity Analysis 31(1), 15-26.

Hynninen, S.-M., Kangasharju, A. & Pehkonen, J. (2009) Matching Inefficiencies, Regional Disparities and Unemployment. Labour 23(3), 481–506.

Håkansson, C. (2014). A divided labour market – on matching on the Swedish labour market after the economic crisis. Sveriges Riksbank Economic Review 2014:2.

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

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.

Lindvall, J. (2012). 'Politics and policies in two economic crises: the Nordic countries' in N. Bermeo & J. Pontusson (eds.), Coping with Crisis: Government Reactions to the Great Recession, New York, NY: Russell Sage Foundation.

Marthin, G. (2012). Measuring Mismatch in the Swedish Labor Market in *Studier i Finanspolitik 2012:3*. Swedish Fiscal Policy Council.

Mortensen, D. T. & Pissarides, C. A. (1994). Job Creation and Job Destruction in the Theory of Unemployment. Review of Economic studies 61(3), 397–415.

Mossfeldt, M. & Österholm, P. (2011). The persistent labour-market effects of the financial crisis. Applied Economics Letters 18(7), 637-642.

Munich, D., Svejnar, J. & Terrel, K. (1998). 'The worker firm matching in transition: (Why) Are Czechs more successful than others?', William Davidson Institute working paper no.107, University of Michigan Business School.

OECD (2012). Employment Outlook 2012, chapter 1. OECD Publishing, Paris.

OECD (2014). OECD Employment Outlook 2014. OECD Publishing, Paris.

OECD (2015). OECD Economic Surveys: Sweden 2015. OECD Publishing, Paris.

OECD (2017). OECD Economic Surveys: Sweden 2017. OECD Publishing, Paris.

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

SOU (2011). Long Term Survey 2011, SOU 2011:11.

Statskontoret (2015). Högre studiebidrag till arbetslösa ungdomar 2015:3. Bilaga 3 - Metodbilaga.

Stock, J. & Watson, M. (2011). Introduction to Econometrics. 3rd ed. Pearson Education, Boston.

Sveriges Riksbank (2012). Monetary Policy Report, October.

Zakova, P. (2013). Sun spots on the Swedish labour market. European Commision 10(1): 1-10.

Appendix

Table 5. Composition of unemployment stock over the years

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Share LTU	0,32	0,37	0,36	0,31	0,44	0,53	0,52	0,51	0,54	0,55	0,59
Share ALMP	0,35	0,27	0,33	0,35	0,46	0,48	0,48	0,49	0,49	0,50	0,50
Share <25	0,21	0,20	0,21	0,25	0,25	0,26	0,27	0,26	0,24	0,22	0,19
Share >55	0,17	0,18	0,18	0,15	0,16	0,16	0,15	0,15	0,15	0,15	0,15
Share PRIMARY	0,27	0,28	0,29	0,26	0,27	0,28	0,29	0,29	0,30	0,32	0,34
Share SECONDARY	0,51	0,50	0,50	0,54	0,53	0,51	0,51	0,50	0,48	0,46	0,42
Share HIGH	0,22	0,22	0,22	0,20	0,21	0,20	0,20	0,21	0,21	0,22	0,23
Share migrant	0,19	0,22	0,25	0,23	0,25	0,28	0,30	0,33	0,37	0,42	0,48

Notes: Share LTU denotes the share of long-term unemployed in the unemployment stock during a year. Share ALMP denotes the share of unemployed participating in labour market programmes. Share <25 years and share >55 years denote the shares of the age groups in the unemployment stock. Primary, Secondary and High denotes the unemployed by level of educational attainment as a share of the unemployment stock. Migrant denotes the share of foreign-born individuals in the unemployment stock.