International Journal of Business and Economics Research
2021; 10(5): 187-202
http://www.sciencepublishinggroup.com/j/ijber
doi: 10.11648/j.ijber.20211005.14
ISSN: 2328-7543 (Print); ISSN: 2328-756X (Online)

Testing Endogeneity, Convexity and Congestion in a Matching Function: Evidence from Finland

Markku Talonen

Faculty of Social Sciences, University of Helsinki, Helsinki, Finland

Email address:
markku.talonen@helsinki.fi

To cite this article:
Markku Talonen. Testing Endogeneity, Convexity and Congestion in a Matching Function: Evidence from Finland. *International Journal of Business and Economics Research*. Vol. 10, No. 5, 2021, pp. 187-202. doi: 10.11648/j.ijber.20211005.14

Received: September 9, 2021; **Accepted**: October 8, 2021; **Published**: October 21, 2021

Abstract: This study contributes to the literature on the efficiency of regional labor markets using matching function to model labor markets and nonparametric methods DEA and FDH to measure efficiency of those markets. DEA has been the most popular method in empirical studies measuring efficiency for an industry and there is also a literature applying DEA to study the efficiency of labor markets. However, this literature neglects two problems important for consistent estimation of a matching function: the possible endogeneity of inputs and non-convexity of the production set. Endogeneity manifests as correlation between inputs and efficiencies. In this paper, we first analyze whether the inputs of the matching function or unemployed jobseekers and open vacancies are exogenous. As our results do not reject exogeneity, we continue treating these inputs exogenous. Next, we evaluate convexity of production set. Testing convexity is an important prerequisite for the use of DEA, because DEA assumes convexity and supplies consistent efficiencies only when the production set is convex. However, convexity is rarely assessed when DEA is applied. In this paper, we evaluate convexity of the production set of the matching function. We use several tests including ones that are based on recently proposed central limit theorems for moments of DEA and FDH estimators. Out of ten tests performed, six ones reject convexity while four ones do not. The tests leave us with a strong belief in non-convexity, and this directs us to apply FDH instead of DEA in the sequel, when we study congestion of inputs. We find strong congestion of open vacancies concerning Helsinki travel-to-work area for several years. In 2017 the loss of matches due to congestion was more than 20 000, amounting to 2.5% of the labor force in Helsinki region, 0.8% in the whole country. Our research with data on 113 travel-to-work areas and 15 public employment (TE-) offices in 2007–19 in Finland, shows huge differences in labor market situation between regions, especially Helsinki and the rest of the country, calling attention from the decision-makers both in firms and government. Also, our study emphasizes the need to pretest data for exogeneity and convexity before applying DEA.

Keywords: Matching Function, Regional Labor Markets, Efficiency, FDH, DEA, Convexity, Congestion

1. Introduction

1.1. Review of Literature

Starting in the 1970s there is a large literature on matching function (MF), which until the millennium is reviewed by Petrongolo and Pissarides [42] and consists of studies estimating MFs using (panel) regression models and assuming MF being mathematically in the Cobb-Douglas form. After that review, there has appeared a new branch of literature, related to MF, focusing on the measurement of efficiency of the matching process, applying stochastic frontier analysis (SFA) as the method. A selected sample of this literature includes studies by Fahr and Sunde [6], Ibourk et al. [30, 31], Ilmakunnas and Pesola [32], Hynninen [28], Ramirez and Vassiliev [45], Hynninen et al. [29], Hillman [27], Jeruzalski et al. [33], Talonen [61]. Also, another related group of studies has appeared measuring the efficiency of regional labor markets (or public employment agencies) using nonparametric data envelopment analysis (DEA) or related methods. The latter is the reference group for our study, and related studies are reviewed in Table 1.

We will see below that applying MF in a basic form involves using the number of new matches as the output (outcome) variable, and two input variables, unemployed jobseekers, and open vacancies, in stock or flow form, depending on the model. This basic form is often extended adding input variables, e.g., variables describing characteristics of jobseekers, like age and education. And there is a branch of literature studying the influence of active labor market policy (ALMP) on matching: e.g., Boeri and Burda [9], Dauth et al. [21], Lehmann [39].

An extended MF is the starting point in most studies applying SFA. On the contrary, few studies applying DEA follow this practice like Khitri et al. [34] and Sheldon [50]. Khitri et al. [34] extend the basic MF adding personnel resources of the public employment offices (PEOs) in inputs. Sheldon [50] adds the staff of the PEOs as an endogenous input and characteristics of jobseekers as exogenous inputs. As seen, all studies in Table 1 include personnel resources in their analyses. We classify studies as applications of matching function if the variables used in them consist of matches as the dependent variable and both unemployed jobseekers and open vacancies as input variables. According to this there are two studies clearly applying matching functions: Khitri et al. [34] and Sheldon [50], and most the rest of the studies reviewed are not classified as MFs because they do not include open vacancies as an input.

Table 1. *Studies on efficiency of regional labor markets or public employment offices using DEA or related nonparametric methods.*

	Talonen-Tuomaala (1994), [59]	**Talonen (1998), [60]**	**Sheldon (2003), [50]**	**Althin-Behrenz (2004), [2]**	**Althin-Behrenz (2005), [3]**
Research period	1992	1995-96	1997-98	1993	1992-95
Number of periods, data frequency	one year	one year	8 months	one year	5 years
Source country of data	Finland	Finland	Switzerland	Sweden	Sweden
Number of regions	187	181	126	345	253
Method	DEA	DEA	DEA	DEA	DEA
Treating environmental variables		2-stage Tobit	2-stage Tobit	2-stage Tobit	
Matching Function:	no	no	extended	no	no
Personnel as an input	yes	yes	yes	yes	yes
Unemployed as an input	yes	yes	yes	no	no
Vacancies as an input	yes	yes	yes	no	no
Input/Output efficiency	input/output	output	output	input	input
Returns to scale (RTS) assumed	CRS, VRS	CRS, VRS	CRS, VRS	CRS	CRS
Returns to scale (RTS) estimated	IRS, CRS, DRS	IRS, CRS, DRS	IRS*		
Productivity measurement		MI			MI
Mean efficiency	0.91 (VRS)	0.83 (CRS), 0.96 (VRS)	0.81 (CRS)	0.7 (CRS)	0.76 (CRS)
Effect of environmental variables in efficiency:					
Share young (<25)		+			
Share aged (>50)		-			
Share basic education		+/-			
Share LTU (>12 m)		-			
Share ALMP			-		
Share of skilled			+		
Personnel/U					
U				+	
V				+	
V/U		-			
U+V			+		

Table 1. *Continued.*

	Vassiliev et al. (2006), [63]	**Althin-Behrenz-Grosskopf-Mellander (2010), [4]**	**Khitri et al. (2011), [34]**	**Riksrevisionen (2012), [49]**	**Andersson-Månsson-Sund (2014), [5]**
Research period	1998-99	1992–1998	2006-2008	2004-2010	2004-2010
Number of periods, data frequency	12 months	7 years	36 months	7 years	7 years
Source country of data	Switzerland	Sweden	Tunisia	Sweden	Sweden
Number of regions	132	256	82	185	185
Method	DEA	Dynamic DEA	DEA	DEA	Dynamic DDF
Treating environmental variables	2-stage	EWL	SW [55]	jobseeker profile	jobseeker profile
Matching Function:	no	no	extended	no	no
Personnel as an input	yes	yes	yes	yes	yes
Unemployed as an input	yes	yes	yes	yes	yes
Vacancies as an input	no	no	yes	no	no
Input/Output efficiency	output	output	output	output	output
Returns to scale (RTS) assumed	VRS	CRS	VRS	CRS	CRS
Returns to scale (RTS) estimated					
Productivity measurement					
Mean efficiency	0.85 (VRS)	0.49 (CRS)	0.83 (VRS)	0.92 (CRS)	0.92 (CRS)
Effect of environmental variables in efficiency:					
Share young (<25)					
Share aged (>50)					

	Vassiliev et al. (2006), [63]	**Althin-Behrenz-Grosskopf-Mellander (2010), [4]**	**Khitri et al. (2011), [34]**	**Riksrevisionen (2012), [49]**	**Andersson-Månsson-Sund (2014), [5]**
Share basic education					
Share LTU (>12 m)					
Share ALMP					
Share of skilled					
Personnel/U			-		
U					
V					
V/U			+		
U+V			-		

Explanations: * DRS for some large offices; Dynamic DDF= dynamic directional distance function; EWL=expected workload; SW [55]= the two-stage method of Simar and Wilson (2007); returns to scale, RTS: CRS=constant, VRS =variable, IRS=increasing, DRS=decreasing returns to scale; MI=malmquist productivity index; U=unemployed jobseekers, V=open vacancies, V/U=tightness of labor market, U+V measures scale of the labor market.

The way Sheldon [50] integrates labor markets and PEOs using an extended MF is consistent with the marginal role of the PEOs on the labor market: numbers of unemployed jobseekers and open vacancies are mainly determined by the behaviors of firms and inhabitants in each region and exogenous from the point of PEOs. Note, however, that exogeneity of unemployment is valid only concerning large unemployment, including both openly unemployed and participants within any active labor market policy (ALMP) program. Instead, open unemployment is not exogenous but can be influenced by PEOs by helping jobseekers in their search for jobs and placing unemployed jobseekers in ALMP programs. It is large unemployment we assume exogenous and use as an input in our MF.

Note that the five studies with Swedish data are different from other ones in several respects. While three of them includes (stocks or flows of) unemployed job seekers in their analysis, none of them includes open vacancies as an input. However, these studies are innovative in taking account of vacancies and jobseekers indirectly, through “attribute” variables like inverted durations of vacancies and unemployment, Althin and Behrenz [2, 3], and “workload” variables reflecting the expected duration of unemployment like Althin et al. [4], Riksrevisionen [49], Andersson et al. [5]. Three studies mentioned last, are innovative also in dividing outputs into intermediate and final ones, applying dynamic DEA, like Fähre and Grosskopf, [25].

Though all studies in our review apply basic DEA or an extension like dynamic DEA or directional distance function (DDF), the models vary by their orientation and returns to scale (RTS) assumed, for example. Most studies measure efficiency in output direction, whether maximum output has been achieved, when inputs are assumed given. In the opposite direction, input efficiency assumes output given and assesses whether it has been achieved using minimum inputs. While the early Swedish studies measure efficiency in input direction, the most recent ones measure output efficiency. In this study, we measure output efficiency, assuming that inputs, numbers of unemployed jobseekers and open vacancies, are exogenous, and not to be influenced by the PEOs. Instead, PEOs are assumed to ease the functioning of the labor market, i.e., how many and soon jobseekers find jobs and employers employees for their open job vacancies. Consequently, the dependent variable in our MF, outflow from unemployment to employment, is endogenous and can be influenced by the regional agents, especially PEOs. According to Calmfors [13], a traditional rationale for ALMP, the programs of which the PEOs exercise, has been to ease the matching process in the labor market. Our assumption on exogenous inputs and endogenous output is consistent with most of the literature using SFA or DEA for the measurement of efficiency of regional labor markets or PEOs. Note, however, that this may not hold for an extended MF, like one with the staffs of PEOs as an input, because the latter might be endogenous.

The role of RTS is two-fold in these studies. First, to compute efficiencies of regions an assumption concerning RTS is needed. In other words, one must select between different DEA models, and the most relevant choice is between the CCR model by Charnes, Cooper and Rhodes [14] and BCC model by Banker, Charnes and Cooper [7]. The former assumes constant returns to scale (CRS) and measures *overall* efficiency while the latter variable returns to scale (VRS) measuring *technical* efficiency. Overall efficiency is the product of technical and *scale* efficiency; to be overall efficient a decision-making unit (DMU) must be both technically and scale efficient, Cooper et al. [18].

Efficiencies from the two basic models, CCR and BCC are used for different purposes. For instance, a planner of the network of services needs information of the best size of DMUs, and for this she needs both overall and technical efficiencies. Technical efficiencies are more proper for comparing performance of DMUs of various size, in case DMUs cannot decide their sizes, because overall efficiency punishes DMUs for being non-optimal size. In half of the studies reviewed here, only overall efficiency is calculated, while two studies calculate only technical efficiency. Two studies using an extended MF compute both overall and technical efficiencies in purpose to assess the quality of returns to scale for DMUs with various size. Of the latter, Sheldon [50] finds increasing returns to scale (IRS), while Talonen [60], studying overall performance of PEOs finds increasing returns to scale for small PEOs, decreasing for large, and constant for “average” ones, implying that there is a best size for PEOs around the average size.

Recently, two problems when measuring efficiency using DEA have been exposed: endogeneity of inputs and convexity of production sets. These problems, which have been mostly neglected in studies on matching function hitherto, are addressed in this paper. Though all studies reviewed in Table 1, employ DEA in basic or extended form, either of these problems has been addressed in none of them. More generally, in the literature employing DEA as a method, these tests are rarely performed.

1.2. The Purpose and Course of This Study

This study emphasizes assessing and accounting for a possible endogeneity problem when the method DEA is applied. This is the first subject to study here, according to the results of which we continue to further tests. If the inputs appear endogenously determined, we cannot supply consistent results for efficiencies (Orme and Smith [40]), and this will bias our tests for convexity, as the latter will use efficiencies from DEA. The results show that one of the inputs, open vacancies is clearly exogenous. The other input, unemployed job seekers, presents endogeneity: according to the classification of Santin and Sicilia [47] it is a positive low endogenous input. These authors show by simulations that it is a positive high endogenous input that can distort the efficiencies and calls for instrumental estimation. We conclude that a positive low endogenous input is not harmful to our analysis, and do not switch to instrumental estimation but continue analysis with the original values of this input.[1]

An important prerequisite for DEA is convexity of the production set, as it is well-known that DEA efficiency measures are not consistent whenever convexity does not hold, see Kneip, Simar and Wilson [38]. In none of the studies presented in Table 1, convexity is evaluated, and this paper contributes to the rare literature of testing convexity of production sets in the context of DEA. Our tests for convexity provide mixed results but leave us with a strong belief in non-convexity. In consequence, we direct to apply FDH instead of DEA in the sequel, when studying congestion of inputs.

The problem of endogeneity in a matching function has been addressed before in some papers, like Borowczyk-Martins et al. [11], but not in the framework of DEA. To the best of our knowledge, this is the first study in the literature on matching function, to assess endogeneity and congestion of inputs and test convexity.

The structure of the paper is as follows. In the next subsection the matching function is presented. Chapter 2 reviews methods DEA and FDH for the measurement of efficiency. Chapter 3 presents data, Chapter 4 the results and Chapter 5 concludes. The tests used in this paper are presented in the Appendices.

1.3. Matching Function

On the labor market, employers are searching for workers, while unemployed jobseekers are searching for job vacancies. There are many channels used to mediate information needed in this process, like advertisements in the media, public and private employment agencies, and private networks. Matching function is an extremely fashionable way to summarize the recruiting processes and their results on the labor market (Blanchard and Diamond [8], Pissarides [43]). According to the idea of MF, "this complicated exchange process is summarized by a well-behaved function that gives the number of jobs formed at any moment in time in terms of the number of workers looking for jobs, the number of firms looking for workers, and a small number of other variables" [42]. In the simplest form, a MF can be written as follows:

$$H_t=f(V_t,U_t) \quad (1)$$

As the relevant period in this study is a year, we note that in (1) H_t denotes all new hirings within year t, V_t job vacancies that are open within year t, while U_t all registered unemployed job seekers within year t. More particularly, U_t denotes the sum of job seekers that were registered unemployed at the beginning of the year t, and those job seekers whose unemployment started within the same year. Similarly, V_t is the sum of job vacancies that were open at the beginning of the year t, and vacancies announced open within the same year.

There are two competing theories for matching: *random (stock-stock)* and *stock-flow*. According to the former, new hirings come into being after the stocks of jobseekers and open vacancies meet and the probability of a match is the probability that an unemployed jobseeker and an open job vacancy meet, multiplied by the probability that this pair is compatible. Using cumulative yearly data, we do not distinguish between random or stock-flow matching. While using monthly data e.g. with stocks of unemployed job seekers and open vacancies at the beginning of each month, and inflows of job seekers and vacancies within each month, we will face the problem of time-aggregation: the stocks measured at the beginning of each month do not measure exactly the real stocks within a month, which vary from day to day, as outflows reduce stocks while inflows increase those. The time-aggregation problem was solved by Gregg and Petrongolo [26] and Coles and Petrongolo [16]. In the following we use data from 2007–19. Our input data consists of stocks and flows as follows. The stock on the first of January amounts to 27% on average of the whole number of unemployed within a year, while the share for the inflow is 73%. The stock at the beginning of each year is not meant to stand for the stock of unemployed job seekers for the full year but stands for those jobseekers whose spell of unemployment has started before the first of January each year, and no matter whether such spells will end or not within the same year. Further, we can imagine that within a year the pool of unemployed changes continuously from hour to hour when some unemployed jobseekers flow in and some other ones out. This process creates virtual stocks, consisting of individual stocks the length of which varies from one day to several weeks and for some job seekers more than a year. However, most jobseekers who start their spells of unemployment within a calendar year also end those within the same year. In summary, using yearly data, we do not see a similar problem of time-aggregation as with monthly or quarterly data (Gregg and Petrongolo [26], Coles and Petrongolo [16]). All inflows and outflows within a year are seen and counted, while numerous stocks (except stocks at the beginning of the year) remain in a black box. On the vacancies side, the

1 To deal with endogeneity of inputs, Santin and Sicilia [47] present a method which is analogous to the method of two-stage least squares.

stocks at the beginning of a year represent 8.4% while inflows of vacancies 91.6% of the total number of vacancies within a year. In summary, our model, and data stand for implicit stock-flow matching.

2. Methods FDH and DEA for the Measurement of Efficiency

In this section, two nonparametric methods for the measurement of efficiency, used here for testing convexity of the production possibilities set, PPS are presented. We do not present all the basic assumptions necessary for constructing PPS for FDH and DEA, but refer to Cooper et al. [18] or Sickles and Zelenyuk [52] for an axiomatic presentation. We focus on assumptions concerning disposability of inputs and outputs and convexity of the PPS. The PPS denoted Ψ is defined as follows:

$$\Psi(x,y)=\left\{(x,y)\middle|x \text{ can produce } y\right\} \tag{2}$$

The method called "free disposal hull", FDH, was introduced by Deprins, Simar and Tulkens [23]. The name of this method is derived from the way the PPS is constructed: it is the hull of the observations when inputs and outputs are assumed freely disposable. The PPS for FDH is the following

$$\hat{\psi}_{FDH}(S_i)=\bigcup_{(X_i,Y_i)\in S_i}\{(x,y)\in R^{p+q}\,|\,y\le Y_i, x\ge X_i\}, \tag{3}$$

where S_i denotes observations consisting of pairs of vectors (x_i,y_i) of inputs $x_i=(x_{1i},x_{2i},\ldots,x_{qi})^T$, and Y_i of outputs $y=(y_{1i},y_{2i},\ldots,y_{pi})^T$ for decision making unit (DMU) indexed i; subscript T denotes the transpose of a vector. It is seen that the PPS is based on all observed inputs and outputs X_i and Y_i added by such input vectors that are larger than or equal to the observed inputs of DMU_i, $x\geq X_i$ and output vectors that are smaller than or equal to the output vector of DMU_i, $y\leq Y_i$. This is the assumption of *free disposal of inputs and outputs*; in other words, it is always possible to increase inputs without decreasing outputs and decrease output without increasing inputs. When the PPS is defined like (3), efficiency of DMU_i ($\lambda(x_i,y_i)$) can be measured in the output increasing direction (output efficiency) as presented by (4):

$$[\lambda(x_i,y_i)]^{-1}=\sup\{\lambda>0\,|\,(x_i,\lambda y_i)\in\hat{\psi}_{FDH}\}, \tag{4}$$

which, in practice can be calculated as follows

$$[\hat{\lambda}_{FDH}(x_i,y_i)]^{-1}=\max_{j|x_j\le x_i}\{\min_{k=1,\ldots,q}\frac{y_j^k}{y_i^k}\}, \tag{5}$$

where i refers to the DMU under measurement, while j refers to the DMUs in the reference set, and k to a component of an output vector, Simar and Wilson [57].

Data envelopment analysis, DEA was introduced by Charnes, Cooper and Rhodes [14], and developed further by Banker, Charnes and Cooper [7] among others. This method has been extremely popular in empirical studies (Emrouznejad [24]). Like FDH, DEA imposes free disposability of inputs and outputs. Additionally, DEA imposes convexity in two different forms. First, in the original model of Charnes, Cooper and Rhodes [14], called the CCR model, production possibilities are constructed as the convex cone of observed inputs and outputs. Second, the model of Banker, Charnes and Cooper [7], called the BCC model, constructs convex hull of observed inputs and outputs for production possibilities. These two models differ by the returns to scale involved: CCR assumes CRS while BCC is consistent with VRS. In CCR, constant returns to scale implies an observed input-output vector (x_i,y_i) can be multiplied by an integer k, or $k(x_i,y_i)=(kx_i,ky_i)$ is possible. When all observed input (output) vectors can be multiplied and added, it results a convex cone (Sickles and Zelenyuk, [52]). Instead, in the BCC model, production possibilities are constructed as convex combinations of observed inputs and outputs, i.e. if x_{i1}, y_{i1} and x_{i2}, y_{i2} are observed, the following linear combination: $a*(x_{i1}, y_{i1})+(1-a)*(x_{i2}, y_{i2})$ is also possible, and the combination is convex because the sum of the coefficients is restricted to one: a+1-a=1.

The BCC model is used in our tests, the PPS of which is presented by (6):

$$\hat{\psi}_{DEA\_VRS}(S_i)=$$

$$\{(x,y)\in R^{p+q}\,|\,y\in Yw, x\in Xw, i^Tw=1, w\in R_+^n\} \tag{6}$$

where subscript DEA_VRS refers to a DEA model assuming variable returns to scale (the BCC model), w is an n-dimensional vector of intensity variables, $w=(w_1,w_2,\ldots,w_n)$, needed to create convex combinations as virtual observations, the condition $i^Tw=1$ is to ensure convexity of those combinations, and i is a n-dimensional vector of ones, $i=(1,1,\ldots,1)^T$.[2] When a PPS is like (6), output efficiency is measured as follows:

$$[\hat{\lambda}_{DEA\_VRS}(x_i,y_i)]^{-1}=\sup_{\lambda,w}\{\lambda\,|\,\lambda y_i\le Yw, x_i\ge Xw, i_n^Tw=1, w\in R_+^n\} \tag{7}$$

Here we measure technical efficiency. The tests for convexity presented below are defined for technical efficiencies, so efficiencies from (5) and (7) are used in our tests [57].

3. Data

The DMUs of our data consist of travel-to-work areas or TTWAs and employment and economic development offices or TE -offices. The latter ones are successors of public employment offices. Our main interest is to study TE -offices but we enlarge the sample by considering TTWAs because our sample increases remarkably enhancing the power of our tests. If we analyzed TE -offices only, we would have 195 DMUs but with TTWAs the number will rise to 1582.

A TTWA is formed by a central municipality and surrounding municipalities from which at least 10 per cent of the labor force commute to the central municipality. A TTWA is named after the

2 Note the dimensions of vectors and matrices: y_i: px1, x_i: qx1, Y: nxp; X: nxq; i=nx1; w=nx1.

central municipality. In this paper, we use TTWAs defined for the year 2019 by Official Statistics of Finland (OSF, [41]). So, there were 37 TTWAs with two or more municipalities and 97 municipalities not belonging to any TTWA. The size of TTWAs varies substantially: the largest area of Helsinki includes 26 municipalities, while the smallest include only 2, and the average is 7 municipalities per TTWA. As we have 37+97=134 DMUs, for which we have data for 13 years, we have 1742 observations available. However, after removing observations with zero values for inputs or outputs for some year, we have a pooled data of 1582 observations.

The dependent variable is planned to measure all new hires on each area. When this information is not available in the statistical data we use, a proxy variable must be used. Three options, often used in empirical studies are: 1) outflow from unemployment into employment, 2) outflow from unemployment, 3) outflow of vacancies or filled vacancies, Broersma, and van Ours [12]. The first alternative is chosen here, as this choice is consistent with the corresponding input variable: unemployed job seekers. The data on our input variables and output variable are taken from the Employment Service Statistics, a part of the Official Statistics of Finland [41]. The dependent variable is a flow variable, which we count cumulatively from the monthly time series. Our first input is the number of unemployed, which is measured as the stock at the beginning of each added by the unemployment spells that started within that year. We note that our measure for unemployment differs from the number of individual jobseekers within a year, as there are possibly multiple spells of unemployment for any jobseeker. Further, we note that with this variable, we measure large unemployment, taking account of both openly unemployed and participants of any ALMP program as well. This follows from the fact that all participants of such programs must have been unemployed for a certain time, typically half a year before entering the program. So, all participants have both a spell of a program and at least one spell of unemployment, the one before the program started. If a participant finds a job after finishing the program, the outflow from (large) unemployment into employment will increase by one, when we assume that without participating this jobseeker would not have found a job. If a jobseeker will not find a job after finishing a program, but will be unemployed again, (s)he will have two spells of unemployment instead of one within a year.[3]

Table 2. *Basic statistics for output and input variables, vacancies-unemployed ratio, and job-finding-rate by TTWAs 2007–2019, N=1582. Source: OSF.*

	Tojob	U_tot	V_tot	V/U	Jobf_rate
minimum	1	106	5	0.007	0.004
mean	2626	9153	3991	0.328	0.310
maximum	68414	261094	270161	1.780	1.000
sd	6540	25361	16153	0.233	0.145
var_coef	2.49	2.77	4.05	0.71	0.47

3 We can conclude that ALMP is successful if increasing the number of participants is associated with more matches and less unemployment spells.

Table 3. *Basic statistics of output and input variables, vacancies-unemployed-ratio, and job-finding-rates by TE-offices 2007–2019, N=195. Source: OSF.*

	Tojob	U_tot	V_tot	V/U	Jobf_rate
minimum	3642	11434	4828	0.22	0.20
mean	21856	60767	36305	0.53	0.38
maximum	66137	221658	313972	1.56	0.74
sd	13041	43847	46846	0.23	0.10
var_coef	0.60	0.72	1.29	0.43	0.26

Tojob=outflow from unemployment into employment (matches); U_tot=number of unemployment spells registered at TE-offices within a calendar year; V_tot=number of open vacancies announced to TE-offices within a calendar year; V/U==V_tot/U_tot; Jobf_rate=job-finding rate=Tojob/U_tot.

The basic statistics for our output and input variables are presented in tables 2 and 3. It is noted that variation is a lot stronger when we look at TTWAs rather than TE-offices.

Figure 1 presents the time patterns of output and input variables for the whole country in 2007–19. Our research period can clearly be divided into two different subperiods: before and after 2011. Though the former includes high unemployment in consequence of small and negative economic growth 2008–09 due to monetary crisis, matches were on a remarkably higher level than 2011 and on. In 2011–2012 matches drop with 36 % and stay on a slightly descending curve in 2013–19, though the number of open vacancies and the tightness of labor market (V/U) are steeply rising in 2016–2019. We note that in 2019 there were 300 000 open vacancies more than 2015, still there were 60 000 matches less. The functioning of the Finnish labor market seems to have deteriorated remarkably.

This is in line with Talonen's [62] results on the efficiency of matching: the efficiency of labor market in 2017–19 was 13% lower than 2007–12 on average.[4] A similar comparison of the volume of ALMP shows that the latter was halved in the latter period compared to the former one.

Figure 2 shows remarkable differences in regional labor markets. The situation in the TE -office of Uusimaa (the first column) – which is identical to the TTWA of Helsinki – deviates most of the rest, with its highest V/U-ratio, lowest job-finding rate and lower than average unemployment rate. A V/U-ratio higher than average seems to be associated with an average u-rate but average or lower job-finding rate. The highest job-finding rate is associated with an average V/U-ratio and a low u-rate.

The low performance of Helsinki TTWA can also be seen from the fact that this region counts 15 % of all matches in Finland though its shares of the unemployed and open vacancies are 22% and 36%, respectively. It is to be noted that the first five TE-offices counted from the left in Figure 2, lie in southern Finland and count 46%, 52% and 64% of matches, unemployed and open vacancies in Finland, respectively.

4 The efficiency is measured as ratio observed/potential matches (outflow from unemployment into employment) where the potential number is calculated as the maximum number of matches related to the numbers of open vacancies and unemployed jobseekers. The potential numbers of matches are calculated using stochastic frontier analysis.

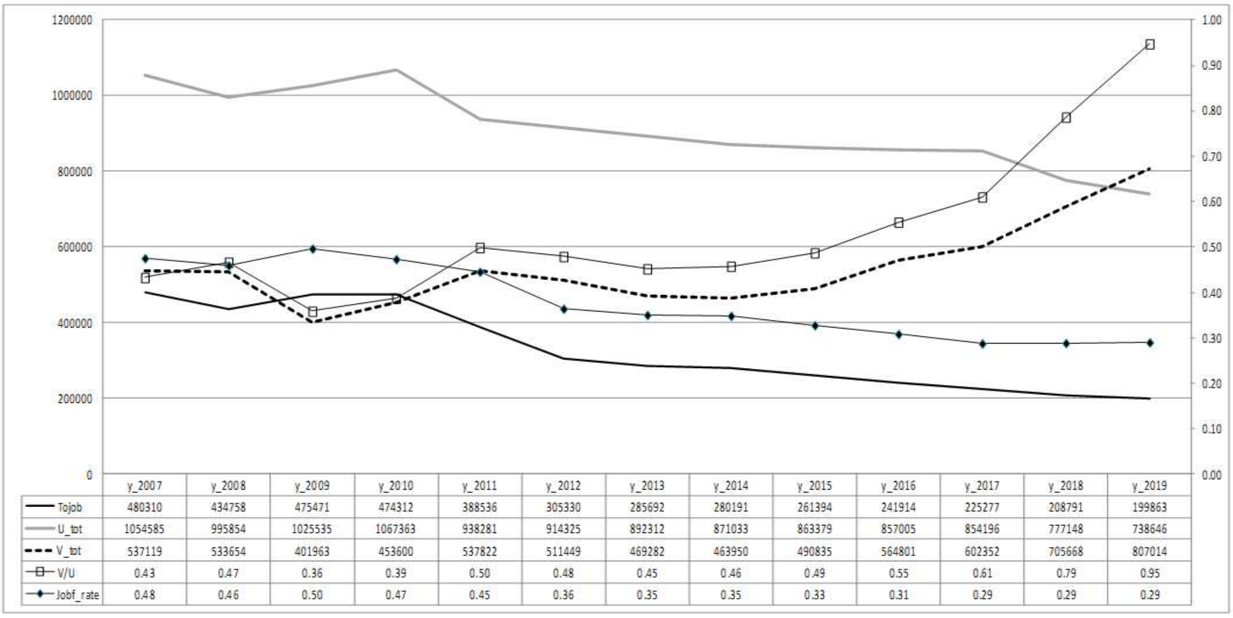

Figure 1. *Output and input variables as cumulative sums for each year; V/U-ratios and job-finding rates as yearly averages in Finland 2007–2019. (V/U-ratio and Job-finding rate on the right axis). Source: OSF. See the explanations for Table 3.*

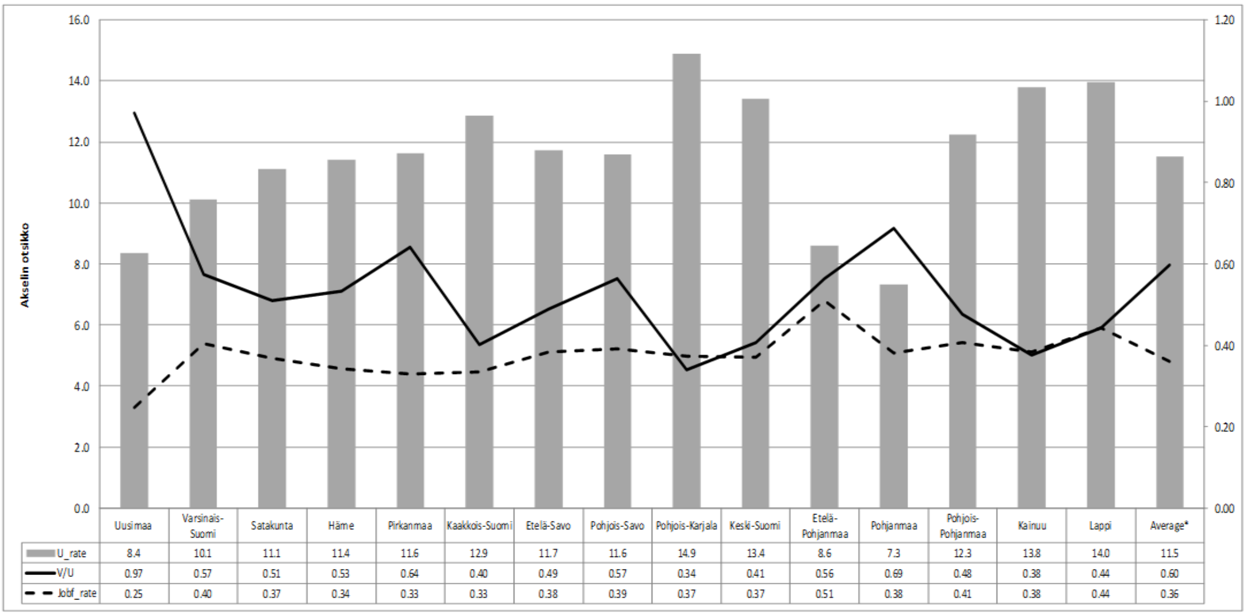

Figure 2. *Rate of unemployment (U_rate), labor market tightness (V/U-ratio) and job-finding rate (Jobf_rate=matches/unemployed) by TE-offices on average 2007–19. Source: OSF.*

Table 4. *Endogeneity or exogeneity of input variables, unemployed jobseekers (U_tot) and open vacancies (V_tot) in TE-offices (N=210)[1] and TTWAs (N=1562) according to the method of Santin and Sicilia [47].*

	γ^*_k	Endogeneity/exogeneity
N=210		
Unemployed (U_tot)	0.063	exogenous
Vacancies (V_tot)	0.002	exogenous
N=1562		
Unemployed (U_tot)	0.29	positive low endogenous
Vacancies (V_tot)	0.19	exogenous

[1] Note that we have 210 observations for TE-offices here, instead 195 ones earlier. When performing this test, data for the year 2020 were available.

4. Results

4.1. Assessing Endogeneity of Inputs in the Basic Matching Function

Endogeneity was assessed using the method of Santin and Sicilia [47], described in Appendix 1. One of our input variables, open vacancies appear exogenous when assessed in both our smaller sample for TE-offices and bigger for TTWAs, see Table 4. The other input, unemployment spells, appear as a positive low endogenous input in our bigger sample. According to Santin and Sicilia [47], a positive medium or high endogenous input can be detrimental to an analysis of efficiency. As the input unemployed jobseekers here is positive low endogenous, we conclude that we can continue analysis without resorting to instrumental estimation.

4.2. Tests for Convexity

While DEA was introduced by Charnes, Cooper and Rhodes [14] presenting the CCR model and Banker, Charnes and Cooper [7] introducing the BCC model, the non-convex alternative FDH was presented around simultaneously by Deprins et al. [23]. From the beginning to now, DEA has been overwhelmingly more popular than FDH in empirical applications. So, convexity of production technology has been accepted, often implicitly. However, the assumption of convexity is not without problems. Cherchye et al. [15] state that convexity assumes away indivisible inputs and outputs, economies of scale and economies of specialization, concluding that there is no good reason for considering convexity of production sets or convexity of input/output sets as generally realistic axioms. They also find that proper tests for convexity are lacking.

Briec et al. [10] seem to be the first to refer to testing convexity. They compared efficiencies calculated assuming convexity or not, assessing the influence of assuming convexity on the level of efficiency under different assumptions on RTS. They compared efficiencies from FDH and DEA calculated assuming similar RTS. They presented an empirical example but did not propose any well-defined procedure for testing convexity.

Simar and Wilson [56] suggested two test statistics for testing convexity of production technology. The first test computes how much the ratio $\hat{\lambda}_{DEA\_VRS} / \hat{\lambda}_{FDH}$ differs from unity, while the other computes the difference between production frontiers constructed using DEA and FDH. The latter test is applied here and explained in detail in Appendix A2.

A new generation of tests for various issues including convexity was recently proposed by Kneip, Simar and Wilson [38] based on new central limit theorems for moments of FDH and DEA estimators by Kneip, Simar and Wilson [37]. These tests are presented in appendices A3.1 and A3.2.

Table 5. *Tests for convexity, pooled data of 37 TTWAs and 97 municipalities in 2007–2019, n=1582.*

Splits&inference	**Test 1a**	**Test 2a**	**Test 3a**	**Test 4a**	**Test 5a**
	SW (2010) [56]	**KSW (2016) [38]**	**KSW (2016) [38]**	**SW (2020) [58]**	**SW (2020) [58]**
	no split & bootstrap	**one split & asymptotic**	**one split & bootstrap**	**multiple splits & bootstrap**	**multiple splits & bootstrap**
m for m-bootstrap	650				
No of splits for n1 (DEA) and n2 (FDH)	0	1	1	10	10
No of splits for bias K	.	1	100	100	100
Tau_5		17.69	17.69	11.06	
K_n					0.30
Tau_2	$1434.7x10^6$				
$p(T \geq T_0)$	0.0165	0.000	0.997	0.000	0.177
confidence intervals					
lower 2.5%	$369.9x10^6$	-0.03	19.92	2.40	0.10
upper 97.5%	$1354.2x10^6$	0.01	36.40	6.78	0.43
Zero hypothesis of convexity:	rejected	rejected	not rejected	rejected	not rejected

Table 6. *Tests for convexity, pooled data of 15 TE-offices in 2007–2019, n=195.*

Splits, inference	**Test 1b**	**Test 2b**	**Test 3b**	**Test 4b**	**Test 5b**
	SW (2010), [56]	**KSW (2016) [38]**	**KSW (2016) [38]**	**SW (2020), [58]**	**SW (2020), [58]**
	no split & bootstrap	**one-split & asymptotic**	**one-split & bootstrap**	**multiple splits & bootstrap**	**multiple splits & bootstrap**
m for m-bootstrap	m=70				
No. of splits S for n1 (DEA) and n2 (FDH), $n_1=n_2=97$	0	1	1	20	20
No. splits for bias K	.	1	100	100	100
Tau_2	$2984x10^6$				
Tau_5		3.37	3.37	4.76	
K_n					0.3
$p(T>T_0)$	0.000	0.000	0.482	0.000	0.046
Confidence intervals					
lower 2.5%	$19.8x10^6$	-0.051	1.09	2.68	0.1
upper 97.5%	$90.3x10^6$	0.051	5.64	3.76	0.35
Zero hypothesis of convexity:	rejected	not rejected	not rejected	rejected	rejected

Explanations for Tables 5 and 6: SW=Simar and Wilson, KSW=Kneip, Simar and Wilson; Tau_2 refers to (9) in Appendix A2; tau_5 to (24) in A3.1; K_n refers to K_n in Step 2 in A4; $p(T>T_0)$ is the probability that the value of the test variable is larger than the value calculated for the test. Note that tests denoted a and b like 1a and 1b etc. are similar but the sample varies from TTWAs denoted with a while TE-offices with b.

All tests applied here, use comparing efficiencies or production frontiers based on one hand on DEA estimator and on the other hand FDH estimator, with a different relation to convexity. In DEA, the maintained hypotheses are convexity and free disposal of inputs and outputs. FDH maintains free disposability but drops convexity. FDH is a consistent estimator of efficiency whether PPS is convex or not. Instead, DEA is consistent only if PPS is convex. This opens a way for testing convexity by calculating efficiencies using both methods, and then comparing whether efficiencies from FDH and DEA are close to each other, which supports convexity of the PPS, or far away, referring to non-convexity of the PPS [38].

The tests for convexity show mixed results. If we take a look at the results concerning our bigger sample (Table 5), we find that three tests reject convexity, while two ones do not. For the smaller sample (Table 6), also three tests reject convexity and two do not, but the results are not consistent for our smaller and bigger sample. In summary, as convexity is rejected in six out of ten in our tests, we are left with a strong suspect that the PPS of the basic matching function is non-convex. To avoid the possible bias due to assuming convexity, though it might not stand, we abandon DEA and continue analysis with FDH.

4.3. Measuring Congestion in the Efficiency of Matching Using FDH

In consequence of the results of the last section, in this section, we abandon DEA and perform an analysis of congestion applying FDH.

For a matching function like (1) it is usually assumed that the marginal products are non-negative: $\partial H/\partial V \geq 0$, $\partial H/\partial U \geq 0$, or the MF is weakly monotone in inputs. This means an increase in vacancies or unemployed may increase or leave intact matches but never decrease them. In case there is congestion, weak monotonicity does not hold, marginal products are negative and increasing vacancies or unemployed job seekers will result in less matches. To be more formal, we cite the definition of congestion in Cooper et al. [17]: "Evidence of congestion is present when reductions in one or more inputs can be associated with increases in one or more outputs - or, proceeding in reverse, when increases in one or more inputs can be associated with decreases in one or more outputs – without worsening any other input or output."

There are several articles on how to measure congestion in DEA models, see the review of Khodabakhshi et al. [35]. However, for FDH models there were no studies until Abbasi et al. [1] proposed a model to assess congestion. The algorithm of Abbasi et al. [1] is used in this study to assess congestion in the production possibilities set constructed as FDH models. This method is sketched in Appendix A5.

The results of our analysis show congestion in the TTWA of Helsinki for several years in 2010–19. The results for this period are presented in Table 7. Note that congestion was analysed also for years 2007–09, but as there were no signs of congestion, these years are dropped from Table 7.

Table 7. *Congestion in Helsinki TTWA 2010–19. Source: author's own calculations.*

Year	TTWA	M_obs	Delta_Y	Delta_U	Delta_V	Strong/Weak	U_tot	V_tot	V/U	V/U_cor
2010	Helsinki	66137					221196	143024	0.65	1.58
2011	Helsinki	54026	904	15153	6185		200675	185375	0.92	2.05
2012	Helsinki	43050	8680	62	202	Strong	200737	185577	0.92	1.91
2013	Helsinki	39507					199436	174048	0.87	1.78
2014	Helsinki	44162					199063	171197	0.86	1.95
2015	Helsinki	45346	8680	3019	2018	Strong	203694	187393	0.92	2.11
2016	Helsinki	44321	1025	7628	38194	Strong	211322	225587	1.07	2.31
2017	Helsinki	46280	21267	462	34966	Strong	221658	224126	1.01	2.21
2018	Helsinki	46280	13745	4534	77912	Strong	205209	267072	1.30	2.77
2019	Helsinki	45784	14241	41	124812	Strong	200716	313972	1.56	3.59

Explanations: M_obs=number of observed matches; Delta_Y=loss of matches due to congestion; Delta_U=congestion in unemployed jobseekers; Delta_V=congestion in open vacancies; Strong /Weak= strong (weak) congestion; U_tot=number of unemployment spells within a year; V_tot=number of open vacancies within a year; V/U=labor market tightness; V/U_cor=labor market tightness corrected with market share in Helsinki.

It is easy to imagine situations, where congestion might appear in matching. Standard MF theory presents a case, where matches are a function of labor market tightness measured by the ratio vacancies/unemployed or V/U-ratio. If CRS is assumed, equation $M=AU^aV^{1-a}$ can be changed into $M/U=A(V/U)^b$, which describes the job-finding-rate as a function of the tightness of the labor market. We can illustrate congestion using this simple model. It is well-known that V/U-ratio varies procyclically. The behaviour of V/U-ratio over the cycle can be described as follows. When economy is in a recession, the V/U-ratio is relatively low, there are a lot of jobseekers competing for few job vacancies, and matching works well, but the number of matches is low because there are not many vacancies open. When a boom starts, there will be more vacancies and the number of matches will rise because there are still many jobseekers available. When the V/U-ratio keeps rising and eventually reaches a high and top level within a cycle, matching gets more difficult and matches do not response easily to the inflow of new vacancies any more, as the pool of unemployed has decreased. Eventually the labor market may end up a situation with congested vacancies when further increase in vacancies will impact less matches: when tightness of the labor market rises, the average time in filling vacancies increases, and there will be less matches in a month or year. In addition, part of open vacancies cannot be filled at all. Scarcity of employable jobseekers is the other side of the coin: when tightness of labor market is very high, “good” workers have been screened out of the pool of jobseekers, and the remaining pool is overpresented by jobseekers that are difficult to employ, like long-term unemployed, aged and less educated people.

Table 7 shows that congestion in Helsinki TTWA appeared in seven years out of ten in 2010–19. The largest losses in matches appeared in 2017–19, when the tightness of labor market exceeded 0.92. It is well-known that in Finland most unemployed jobseekers are registered at PEOs while around half of the open vacancies are announced at PEOs. We also calculated tightnesses of labor market correcting the number of open vacancies in Helsinki TTWA multiplying the latter by the reciprocal of the share of all vacancies announced to PEOs (“market share”) in Helsinki region. Information on the market share is based on employer interviews carried out by Statistics Finland and published by the Ministry of Economic Affairs and Employment in yearly reports [48]. The corrected labor market tightnesses for Helsinki TTWA exceed 2 in 2015–18 and reach 3.6 in 2019. The interviews mentioned above also report information of difficulties in recruitment and reasons for these. According to reports in 2018 and 2019 the share of firms suffering from the shortage of labor were 19 and 22 per cent respectively, while the lossess of matches due to labor shortage in Helsinki TTWA were 18 000 and 25 600, respectively. The first of these is roughly in line with our estimates of losses of matches due to congestion, but the second nearly doubles our estimate for congestion. Sticking to our estimates, we can calculate that without congestion, the rate of unemployment in Helsinki TTWA would have been 1.6 percentage points lower than the observed one, and the influence on the average rate of unemployment in the whole Finland would be 0.5 percentage units.

5. Conclusion

This study contributes to the literature on the efficiency of regional labor markets or public employment offices by several ways. First we study possible endogeneity of inputs in the basic matchng function. As we find no serious problem of endogeneity, we continue testing convexity. The results for convexity were mixed: six tests rejected convexity while four ones did not. We are left with a considerable suspect that the PPS of the basic MF is non-convex. In consequence, we abandon DEA and continue doing an analysis of congestion using FDH. Based on these results we emphasize testing endogeneity of inputs and convexity of the PPS as a prerequisite for using DEA in the context of matching function.

Also, we estimate congestion on the Finnish labor markets finding that the travel-to-work area of Helsinki was for many years congested strongly with open vacancies, and somewhat congested with unemployed jobseekers. The top year for congestion was 2017 when there were 31% less matches because of congestion of open vacancies and for the same reason, the rate of unemployment in Helsinki TTWA was 1.6 percentage points higher than if there were no congestion. The congestion in Helsinki TTWA sends a strong message to employers and regional policy makers: to consider alternative locations for investmenets instead Helsinki region.

Appendix

The software needed for the tests in these appendices was written by the author, using r-language and the following packages: Benchmarking, Rsampling and fastmatrix, RDevelopment Core Team [46].

Appendix 1. Assessing Endogeneity of Inputs Using the Method by Santin and Sicilia [47]

Recently, a few studies have paid attention to the possible endogeneity problem when we estimate a matching function. Santin and Sicilia [47] claim that endogeneity is rarely taken account of when DEA is used to measure efficiency in an industry.[5] In production functions, exogeneity holds if inputs are not correlated with efficiencies. Cordero, Santin and Sicilia [20] showed by Monte Carlo experiments that DEA estimates can be severely biased if one input is highly and positively correlated with efficiencies. They also show that the performance of DEA is only deteriorated by a medium or high positive correlation between an input and efficiencies. Santin and Sicilia [47] propose a heuristic method to assess and deal with endogeneity in DEA models. We follow their method in assessing endogeneity of inputs in our data. The method uses bootstrap to draw numerous new pseudo-samples from the original sample, calculates efficiencies for each DMU in each sample, then calculates correlations between each input variable and efficiencies. The bootstrap suggested is naïve, i.e., as many DMUs as in the original sample (n) are drawn with replacement. If some correlation appears positive and high, the respective input is replaced with values calculated as a function of some relevant instrument variables, i.e., the latter are highly correlated with the respective input but zero correlated with efficiencies. The algorithm of Santin and Sicilia [47] is described as follows.

1. Randomly draw a bootstrap sample with replacement $\chi^*=\{(X^*_{ib},Y^*_{ib}),\ i=1,\dots,n\}$ from the empirical dataset $\chi=\{(X_i, Y_i)\ i=1,\dots,n\}$.
2. Compute the efficiency score λ_i for each DMU according to (5) above. Note that we apply output oriented FDH here instead of DEA in [47].
3. For each input k=1, …, p, compute the Pearson correlation coefficient between the estimated efficiency score λ^*_{ib} and the input x^*_{ik}

$$\rho^*_{k,b}=\mathrm{corr}(x^*_{ik}, \lambda^*_i, i=1,\dots,n; k=1,\dots,p).$$

5 Earlier, Borowczyk-Martins et al. [11] refer to endogeneity problem in the estimation of a matching function, proposing that the matching function elasticities are exposed to an endogeneity bias because of simultaneity in determining the job-finding rate, cost of posting a vacancy and labor market tightness. They also address endogeneity but assume constant returns to scale, while we are employing models with variable returns to scale.

4. Repeat steps 1 to 3 B times to obtain a bootstrap set of correlations $\{\rho^*_{kb}, b=1,\ldots,B\}$ for each input.
5. For each input x^*_{ik} compute $\gamma^*_k = \frac{1}{B}\sum_{1}^{B}[I_{[0,1]}(\rho^*)]$, $k=1,\ldots,p$, where I is an indicator function defined by $I_{[0,1]}(\rho^*) = 1$ if $0 \le \rho^*_k \le 1$,=0 otherwise.
6. Classify the inputs according to the proposed heuristic as follows:

If $\gamma^*_k < 0.25$, input k is an exogenous or negative endogenous input.

If $0.25 \le \gamma^*_k \le 0.5$, input k is a positive low endogenous input.

If $0.5 \le \gamma^*_k \le 0.75$, input k is a positive medium endogenous input.

If $0.75 \le \gamma^*_k$, input k is a positive high endogenous input.

Classification in step 6 is of a rule of thumb kind, Santin and Sicilia [47].

Appendix 2. The Test for Convexity by Simar and Wilson [56]

The first test of convexity is by Simar and Wilson (2010) [56], denoted Test 1 here and the tables 5 and 6 in the main text, compares ratio of efficiencies from DEA to efficiencies from FDH for each DMU. If the zero hypothesis of convexity holds, efficiencies from DEA will not be far from efficiencies from FDH, or the corresponding ratio far from one. The first test is the following:

$$\hat{\tau}_1(S_n) = (n^{-1})\sum_{i=1}^{n}\left(\frac{\hat{\lambda}_{FDH}(x_i, y_i \mid S_n)}{\hat{\lambda}_{DEA\_VRS}(x_i, y_i \mid S_n)} - 1\right) \ge 0, \quad (8)$$

where S_n denotes inputs x_i and outputs y_i in the original (full) sample $S_n=(x_i,y_i)$, $(i=1,2,\ldots,n)$. Note that we use here output distance functions, $\hat{\lambda}_i \le 1$, and it is well known that FDH gives at least as high an efficiency as DEA, which means $\hat{\lambda}_{i,FDH} \ge \hat{\lambda}_{i,DEA\_VRS}$ which implies the ratio in the parenthesis of equation (8) ≥ 1, and the test score will be non-negative.

The alternative test of Simar and Wilson [56] that we name Test 2 here, is measuring the mean integrated squared difference between the production frontiers determined by DEA_VRS and FDH:

$$\hat{\tau}_2(S_n) = (n^{-1})\sum_{i=1}^{n} D_{2,i}^T D_{2,i} \ge 0, \quad (9)$$

where $D_{2,i} = (y_i\hat{\Phi}_{i,DEA\_VRS} - y_i\Phi_{i,FDH})$ is a q-vector (q is the number of outputs), and we have adapted the approach of Simar and Wilson [56] with input orientation to output orientation, moving each output observation onto the production frontier. Note that in (9) we use Farrel-type output efficiencies Φ_i which are reciprocals of output distance functions used elswhere in this paper: $\Phi_i=1/\lambda_i$, so $\Phi_i \ge 1$.

Simar and Wilson [56] also performed Monte Carlo trials to reveal the performance of their two tests. In these tests the latter one or Test 2 performed better for size and power, so we chose that test as one of our tests for convexity of MF. Note that this test differs from later ones in neglecting bias correction, which was made explicit in the tests of Kneip, Simar and Wilson [37, 38]. In practice, we performed this test using m-bootstrap, also proposed by Simar and Wilson [56], the consistency of which was proven by Kneip, Simar and Wilson [36]. This is performed drawing pseudo-samples of size m<n from the original sample. Here all m<n will do but the size of m can be trimmed optimal utilizing the idea of Politis et al. [44]. Using this method provided us with an optimal size m=650; recall that our data consists of 1582 observations. We also did the same test with our smaller data with 195 observations on TE-offices. In this case we find m=70 to be the optimal size for a bootstrap sample.

In our m-bootstrap, 2000 pseudo samples of size 650 (70) were drawn from the original sample of 1582 (195) observations, and test values were calculated according to the following bootstrap principle:

$$m^{\kappa}\sqrt{m}\,\hat{\tau}_2 \xrightarrow{\text{approx}} n^{\kappa}\sqrt{n}\,\tau_2 \quad (10)$$

The zero hypothesis is rejected whenever $n^{\kappa}\sqrt{n}\,\tau_n(S_n) > q_{m,n}(1-\alpha)$, where $q_{m,n}$ is the $(1-\alpha)$ quantile of the bootstrap distribution approximated by

$$\hat{G}_{m,n}(\alpha) = \frac{1}{B}\sum_{b=1}^{B} I(m^{\kappa}\sqrt{m}\,\tau_m(S_m^{*,b}) \le \alpha) \quad (11)$$

where I(.) means the indicator function and $\{\tau_m(S_m^{*,b})\}_{b=1}^{B}$ is the set of B bootstrap estimates.

Appendix 3. Tests for Convexity Proposed by Kneip, Simar and Wilson [38]: Test 2 and Test 3

Appendix 3.1. Test 2 Relying on Asymptotics

These tests are based for comparison of efficiencies provided by FDH and DEA_VRS estimators; recall that the latter is calculated assuming VRS. Under the null hypothesis, both estimators are consistent, but under the alternative hypothesis, only FDH is consistent. For the test, the sample X_n is split into two independent parts, X_{1,n_1} and X_{1,n_2} such that $n_1=[n/2]$, $n_2=n-n_1$, $X_{n_1} \cup X_{n_2} = X_n$ and $X_{n_1} \cap X_{n_2} = \varnothing$, where [a] denotes the integer part of a. The expected values for DEA_VRS and FDH efficiencies are calculated for subsamples 1 and 2, respectively, as follows:

$$\hat{\mu}_{DEA\_VRS,n_1} = n_1^{-1}\sum_{X_i,Y_i \in X_{1,n_1}} \hat{\lambda}_{DEA\_VRS}\left(X_i, Y_i \mid X_{1,n_1}\right) \quad (12)$$

$$\hat{\mu}_{DEA\_VRS,n_2} = n_2^{-1}\sum_{X_i,Y_i \in X_{2,n_2}} \hat{\lambda}_{DEA\_VRS}\left(X_i, Y_i \mid X_{2,n_2}\right) \quad (13)$$

Variances are calculated as follows:

$$\hat{\sigma}^2_{DEA\_VRS,n_1} = n_1^{-1} \sum_{(X_i,Y_i)\in X_{1,n_1}} \{(\hat{\lambda}_{DEA\_VRS}(X_i, Y_i \mid X_{1,n_1}) - \hat{\mu}_{DEA\_VRS,n_1}\}^2 \quad (14)$$

$$\hat{\sigma}^2_{FDH,n_2} = n_2^{-1} \sum_{(X_i,Y_i)\in X_{2,n_2}} \{\hat{\lambda}_{FDH}(X_i, Y_i \mid X_{2,n_2}) - \hat{\mu}_{FDH,n_2}\}^2 \quad (15)$$

The consistency of estimators (12) – (15) is based on Theorem 4.1 in Kneip, Simar and Wilson [37]. To construct estimates for bias, each subsample is split into two mutually exclusive and collectively exhaustive subsamples X_{l,n_l} for $l \in \{1,2\}$, $X^{(1)}_{l,m_l,1,k}$ and $X^{(2)}_{l,m_l,2,k}$. For $j \in \{1,2\}$ we compute:

$$\hat{\mu}^{(j)}_{DEA\_VRS,m_{1,j},k} = m_{1,j}^{-1} \sum_{(X_i,Y_i)\in X^j_{1,m_1}} \hat{\lambda}_{DEA\_VRS}(X_i, Y_i \mid X^{(j)}_{1,m_{1,j},k}) \quad (16)$$

$$\hat{\mu}^{(j)}_{FDH,m_{2,j},k} = m_{2,j}^{-1} \sum_{(X_i,Y_i)\in X^j_{2,m_{2,j},k}} \hat{\lambda}_{FDH}(X_i, Y_i \mid X^{(j)}_{2,m_{2,j},k}) \quad (17)$$

$$\tilde{\mu}^*_{DEA\_VRS,n_1,k} = 0.5(\hat{\mu}^{(1)}_{DEA\_VRS,m_{1,j},k} + \hat{\mu}^{(2)}_{DEA\_VRS,m_{1,j},k}) \quad (18)$$

$$\tilde{\mu}^*_{FDH,n_1,k} = 0.5(\hat{\mu}^{(1)}_{FDH,m_{1,j},k} + \hat{\mu}^{(2)}_{FDH,m_{1,j},k}) \quad (19)$$

$$\tilde{B}_{DEA\_VRS,k_1,n_1,k} = (2^{k_1}-1)^{-1}(\tilde{\mu}^*_{DEA\_VRS,n_1,k} - \hat{\mu}_{DEA\_VRS,n_1}) \quad (20)$$

$$\tilde{B}_{FDH,\kappa_2,n_2,k} = (2^{\kappa_2}-1)^{-1}(\tilde{\mu}^*_{FDH,n_2,k} - \hat{\mu}_{FDH,n_2}) \quad (21)$$

where subscript k refers to the number of a bootstrap round; when there are K rounds, biases can be calculated

$$\hat{B}_{VRS\_DEA,\kappa_1,n_1} = \sum_{k=1}^{K} \tilde{B}_{DEA\_VRS,\kappa_1,n_1,k} \quad (22)$$

$$\hat{B}_{FDH,\kappa_2,n_2} = \sum_{k=1}^{K} \tilde{B}_{FDH,\kappa_2,n_2,k} \quad (23)$$

Finally we have a test score:

$$\hat{\tau}_{5,n} = \frac{(\hat{\mu}_{FDH,n_2} - \hat{\mu}_{DEA\_VRS,n_1})(\hat{B}_{FDH,\kappa_2,n_2} - \hat{B}_{DEA\_VRS,\kappa_1,n_1})}{\sqrt{\frac{\hat{\sigma}^2_{FDH,n_2}}{n_2} + \frac{\hat{\sigma}^2_{DEA\_VRS,n_1}}{n_1}}} \quad (24)$$

$$\xrightarrow{L} N(0,1)$$

Note that in the formulas above the convergence rate for DEA_VRS equals $\kappa_1=2/(p+q+1)$, while the convergence rate of FDH equals $\kappa_2=1/(p+q)$. In our empirical part we have accordingly $\kappa_1=1/2$, and $\kappa_2=1/3$. Note also that (24) can be used in cases, where $p+q\leq3$, like our case where p=2 and q=1, [38].

In the tests above, the original sample is split into two subsamples. This can be done in numerous ways. To make this test replicable, Daraio, Simar and Wilson [22] presented an algorithm for shuffling the original sample before splitting it into subsamples. This algorithm is used in our tests when only one split is used, like tests 2 and 3.

In our one-split tests, 2 and 3 bootstrap is not necessarily needed, as the asymptotics for the tests are known. Kneip, Simar and Wilson [38] propose using bootstrap as an alternative method to asymptotics. Following this, we apply asymptotics for Test 2 and bootstrap for Test 3.

Note finally that these tests using data on TTWAs are published in Table 5, while for the data on TE-offices in Table 6.

Appendix 3.2. A Test of Convexity Based on Bootstrap: Test 3

Our Test 3 is a bootstrap-version of Test 2, the algorithm of which is as follows.

Step 1. The original sample of 1582 TTWAs in data 1 are shuffled using the algorithm of Daraio, Simar and Wilson [22], and divided into two subsamples of sizes 791. The first 791 observations in order in data 1 $\{(X_i, i=1,\ldots,1582\}$ make the first subsample S_1, $\{S_{1,i}, i=1,\ldots,791\}$ and the last 791 the second subsample S_2, $\{S_{2,i}, i=1,\ldots,791\}$. Efficiencies are calculated applying DEA for the first subsample S_1 and FDH for the second subsample applying (7) and (4), respectively. The mean efficiencies for the two subsamples, $\hat{\mu}_{DEA\_VRS,n_1}$ and $\hat{\mu}_{FDH,n_2}$ where $n_1=n_2=791$, are calculated using (12) and (13), and the respective variances using (14) and (15): $\hat{\sigma}^2_{DEA\_VRS,n_1}$, $\hat{\sigma}^2_{FDH,n_2}$.

Step 2. Let $Z_{1,n1} = \{\hat{\lambda}_{DEA\_VRS}(i), i=1,\ldots,n_1\}$ and $Z_{2,n2}=\{\hat{\lambda}_{FDH}(i), i=1,\ldots,n_2\}$.

Draw n_1 observations uniformly and independently with replacement from S_1 to make bootstrap sample S_1^* and similarly from S_2 for S_2^*. Note that a naïve resampling works here, Kneip, Simar and Wilson [38].

Step 3. Shuffle subsamples S_1 and S_2 and divide further into two subsamples $S_1\_m_1$, $S_1\_m_2$, $S_2\_m_1$ and $S_2\_m_2$ by taking first and last 395 in the order of each subsample. Then calculate the biases for DEA and FDH, $\tilde{B}_{DEA\_VRS,\kappa_1,n_1,k}$ and $\tilde{B}_{FDH,\kappa_2,n_2,k}$ using (20)-(21). Do this K times ending up with $\hat{B}_{VRS\_DEA,\kappa_1,n_1}$ in (22) and similarly $\hat{B}_{FDH,\kappa_2,n_2}$ in (23). Calculate $\hat{\tau}_{5,n}$ according to (24):

$$\hat{\tau}_{5,n} = \frac{(\hat{\mu}_{FDH,n_2} - \hat{\mu}_{DEA\_VRS,n_1})(\hat{B}_{FDH,\kappa_2,n_2} - \hat{B}_{DEA\_VRS,\kappa_1,n_1})}{\sqrt{\frac{\hat{\sigma}^2_{FDH,n_2}}{n_2} + \frac{\hat{\sigma}^2_{DEA\_VRS,n_1}}{n_1}}}$$

Step 4. Repeat steps 2 and 3 for B-1 times to provide a bootstrap sample $\tilde{B} = \{\hat{\tau}_{5,n_1,n_2,b}\}_{i=1}^{B}$.

Step 5. Calculate a confidence limit at risk level α:

$[\hat{\xi}_{n_1,n_2} - z^*_{1-\alpha/2}\hat{s}_{n_1,n_2}, \hat{\xi}_{n_1,n_2} - z^*_{\alpha/2}\hat{s}_{n_1,n_2}]$, where $\hat{\xi}_{n_1,n_2} = \hat{\mu}_{DEA\_VRS,n_1} - \hat{\mu}_{FDH,n_2}$, $\hat{s}_{n_1,n_2}$ is the denominator in (24), $z^*_{\alpha/2}$ and $z^*_{1-\alpha/2}$ are taken from the bootstrap distribution.

Step 5. If the confidence limit from the previous step does not include 0, reject the zero hypothesis, otherwise fail to reject it.

The five steps above were described using our bigger sample of TTWAs. The same algorithm was adapted to our smaller sample of TE-offices, too. The results are presented in Table 5 for the sample of TTWAs and Table 6 for TE-offices.

Appendix 4. Testing Convexity Using Bootstrap and Taking Advantage of Multiple Splits of the Original Sample: Test 4 and Test 5

Here we follow the method by Simar and Wilson [58], while adapt it to testing convexity in the *output* oriented measurement of efficiency.

Step 1. The original sample of 1582 TTWAs are shuffled and divided into two subsamples of sizes 791. The first 791 observations in order in data 1 $\{(X_i, i=1,\ldots,1582\}$ make the first subsample S_1, $\{S_{1,i}, i=1,\ldots,791\}$ and the last 791 the second subsample S_2, $\{S_{2,i}, i=1,\ldots,791\}$. Compute $T_{5,n}$ according to (24).

Step 2. Repeat Step 1 s-1 times to obtain $\{\hat{\tau}_{5,n,j}\}_{j=1}^{s}$. Compute $\bar{T}_n = (1/s)\sum_{j=1}^{s}\hat{\tau}_{5,n,j}$. With $\{\hat{\tau}_{5,n,j}\}_{j=1}^{s}$ we also have s corresponding p-values $P = \{\hat{p}_j\}_{j=1}^{s}$ which form an empirical cumulative distribution function $\hat{F}_{s,\hat{p},n}(u) = (1/s)\sum I(\hat{p}_j \le u)$, where I(.) is an indicator function. Compute $\hat{K}_n = \sup_{u\in[0,1]}\left|\hat{F}_{s,\hat{p},n}(u) - u\right|$.

Step 3. Compute $\hat{\lambda}_{DEA\_VRS,i}$ for i=1,…,n (i.e. for the full original sample). Set b=0.

Step 4. Increment b by 1. Draw k_i, i=1,…,n, independently and with replacement from the set of integers 1 through n, such that each integer has probability of 1/n of being selected in a draw, and set $\hat{\lambda}^*_{DEA\_VRS,i} = \hat{\lambda}_{DEA\_VRS,k_i}$.

Step 5. Create a bootstrap sample $X^*_n = \left\{X^*_i, Y^*_i\right\}_{i=1}^{n}$, where $Y^*_i = Y_i\hat{\lambda}_{DEA\_VRS,i}/\hat{\lambda}^*_{DEA\_VRS,i}$ and $X^*_i = X_i$. In words, observations are first moved to the production frontier and then away from it.

Step 6. Analogously to step 1, randomly shuffle the observations in X^*_n and split into two subsamples X^*_{1,n_1} and X^*_{2,n_2} and compute the test statistic $\hat{T}^*_{5,n}$.

Step 7. Repeat Step 6 s times to obtain $T^*_b = \left\{\hat{T}_{5,n,j}\right\}_{j=1}^{s}$. Compute $\hat{T}^*_{n,b} = (1/s)\sum_{j=1}^{s}\hat{T}^*_{5,n,j}$, the set P^*_b of p-values corresponding to the elements of T^*_b and $K^*_{n,b}$ using the values P^*_b as in step 2.

Step 8. Repeat steps 6 and 7 B-1 times to obtain $\left\{\hat{T}^*_{n,b}\right\}_{b=1}^{B}$ and $\left\{\hat{K}^*_{n,b}\right\}_{b=1}^{B}$.

Step 9. Compute

$$\hat{p}_T = \frac{\#\left\{T^*_{n,b} \ge \bar{T}_n\right\}}{B}$$

$$\hat{p}_K = \frac{\#(\hat{K}^*_{n,b} \ge \hat{K}_n)}{B}$$

For a test of size α, reject the null hypothesis if $\hat{p}_T$ or $\hat{p}_K$ are less than α. Note that though described with numbers from the sample of TTWAs, calculations were done similarly for TE-offices.

Note also that the results of test 4 and test 5 for TTWAs as data are presented in Table 5, and for TE-offices as data in Table 6.

Appendix 5. Testing Congestion Using the Method of Abbasi et al. [1]

Algorithm of Abbasi et al. [1]:

Step 1. Calculate the optimal value of the following model:

$$Z_{FDH^{-1}} = \sum_{r=1}^{s}(y_{r,q} - y_{r,p}) = \max_{j\in D_p}\sum_{r=1}^{s}(y_{rj} - y_{rp}) \qquad (25)$$

where q refers to DMUs (q=1,…,q), p to the DMU under calculation and r to outputs (r=1,…,s), and where $D_p=\{j\in 1,\ldots,n|x_j\ge x_p \text{ and } y_j\ge y_p\}$.

Step 2. Let $\hat{y}_p = y_p + s^{+*}$ where $s^{+*} = y_q - y_p$ Obtain the optimal value of the following model

$$Z_{FDH} = \max_{j\in\hat{D}_p}\sum_{r=1}^{s}(y_{rj} - \hat{y}_{rp}) \qquad (26)$$

where $\hat{D}_p = \left\{j\in(1,\ldots,n)\middle| x_j \le x_p \text{ and } y_j \ge y_p\right\}$

Step 3. If $Z_{FDH}>0$, then DMU_j is congested, so go to part b). Furthermore, if there exist $j\in\hat{D}_p$ such that $x_j<x_p$ and $y_j>\hat{y}_p$, then DMU_j is strongly congested. If $Z_{FDH}=0$, then DMU_j is not congested and stop here.

part b)

Step 4. Define K_p as follows:

$$K_p = \left\{j\in\hat{D}_p\middle| Z_{FDH} = \sum_{r=1}^{s}(y_{rj} - \hat{y}_p)\right\} \qquad (27)$$

then calculate

$$\alpha^* = \min_{j \in K_p} \sum_{i=1}^{m} (x_{ip} - x_{ij}) \quad (28)$$

Step 5. Define T_p as follows

$$T_p = \left\{ j \in K_p \middle| \alpha^* = \sum_{i=1}^{m} (x_{ip} - x_{ij}) \right\} \quad (29)$$

For $j \in T_p$ define s_i^{c*} as the amount of congestion in ith input of DMU_p and s_r^{+*} as reduction amount of rth output due to congestion as follows:

$$s_i^{c*} = x_{ip} - x_{ij}, \quad i=1,...,m \quad (30)$$

$$\hat{s}_r^{+*} = y_{rj} - \hat{y}_{rp}, \quad r=1,...,s \quad (31)$$

Abbasi et al. [1].

References

[1] Abbasi M., Jahanshaloo G. R., Rostamy-Malkhlifeh and Lofti H. (2014). Estimation of Congestion in Free Disposal Hull Models Using Data Envelopment Analysis. *The Scientific World Journal*, Volume 2014, Hindawi Publishing Corporation.

[2] Althin, R. and L. Behrenz (2004). An Efficiency Analysis of Swedish Employment Offices, *International Review of Applied Economics*, 18, s. 471–482.

[3] Althin, R. and L. Behrenz (2005). Efficiency and productivity of employment offices: evidence from Sweden, *International Journal of Manpower*, 26, s. 196–206.

[4] Althin, R., Behrenz L. R., Färe R., Grosskopf S. and Mellander E. (2010). Swedish employment offices: A new model for evaluating effectiveness, *European Journal of Operational Research*, 207, s. 1535–1544.

[5] Andersson C., Månsson J. and Sund K. (2014). Technical efficiency of Swedish employment offices, *Socio-Economic Planning Sciences 48, 57–64*.

[6] Fahr R. and Sunde U. (2006). Regional dependencies in job creation: an efficiency analysis for Western Germany, Applied Economics, 38: 10.

[7] Banker, R. D., Charnes A., and Cooper W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Management Science 30, 1078–1092.*

[8] Blanchard O. P., Diamond P. (1989). The Beveridge-curve. *Brookings Papers on Economic Activity*, 1: 1989, 1–76.

[9] Boeri T., Burda M. C. (1996). Active labor market policies, job matching and the Czech miracle. *European Economic Review 40, 805–817*.

[10] Briec W., Kerstens K. and Eeckaut P. (2004). Non-convex Technologies and Cost Functions: Definitions, Duality and Nonparametric Tests of Convexity, *Journal of Economics 81 (2), 155–192*.

[11] Borowczyk M., Jolivet G. and Postel-Vinay F. (2013). Accounting for endogeneity in matching function estimation, *Review of Economic Dynamics*, 2013, vol. 16, issue 3, 440–451.

[12] Broersma L., van Ours J. C. (1999). Job searchers, job matches and the elasticity of matching. *Labour Economics, 6 (1), 77–93*.

[13] Calmfors L. (1994). Active Labour Marekt Policy and Unemployment – A Framework for the Analysis of crucial Design Features, OECD Economic Studies No. 22.

[14] Charnes, A., Cooper W. W. and Rhodes E. (1978). Measuring the efficiency of decision-making units, *European Journal of Operational Research* 2, 429–444.

[15] Cherchye L., Kuosmanen T. and Post T. (2001). Alternative treatments of congestion in DEA: A rejoinder to Cooper, Gu, and Li, *European Journal of Operational Research*, vol. 132, 75–80.

[16] Coles M. G. and Petrongolo B. (2008). A Test Between Stock-Flow Matching and the Random Matching Function Approach, *International Economic Review, 49* (4), *1113–1141.*

[17] Cooper W. W., Deng H., Huang Z. M. and Li S. X. (2002). A one model approach to congestion in data envelopment analysis, *Socio-Economic Planning Sciences*, vol. 36, no. 4, pp. 231–238.

[18] Cooper W. W., Seiford L. M and Tone K. (2007). Data Envelopment Analysis, A Comprehensive Text with Models, Applications, References and DEA-Solver Software, second edition. Springer Science and Business Media.

[19] Cordero, J. M., Santín, D. and Sicilia, G. (2017). Dealing with the Endogeneity Problem in Data Envelopment Analysis. *Expert Systems with Applications* 68, 127–184.

[20] Cordero, J. M, Santín, D, & Sicilia, G. (2015). Testing the accuracy of DEA estimates under endogeneity through a Monte Carlo simulation, *European Journal of Operational Research, 244* (2), 511–518.

[21] Dauth W., Hujer R. and Wolf K. (2010). Macroeconometric Evaluation of Active Labor Market Policies in Austria, The Institute for the Study of Labor, Discussion Paper No. 5217.

[22] Daraio C., Simar L. and Wilson P. W. (2016). Nonparametric Estimation of Efficiency in the Presence of Environmental Variables, University of Rome, Technical Report n. 2

[23] Deprins, D., Simar L. and Tulkens H. (1984). Measuring labor inefficiency in post offices, in M. M. P. Pestieau and H. Tulkens, eds., The Performance of Public Enterprises: Concepts and Measurements, Amsterdam: North-Holland, pp. 243–267.

[24] Emrouznejad, A. (1995–2011). Data Envelopment Analysis Homepage, www.DEAzone.com, last viewed 05.06.2020.

[25] Fähre R. and Grosskopf S. (1996). Productivity and intermediate products: A frontier approach, *Economics Letters 50, 65–70.*

[26] Gregg P. and Petrongolo B. (2005). Stock-Flow Matching and the Performance of the Labor market, *European Economic Review, 49, 1987–2011.*

[27] Hillman K. (2009). Does the Hartz IV Reform have an Effect on Matching Efficiency in Germany? A Stochastic Frontier Approach. Munich Personal RePEc Archive, Paper No. 22295. http://mpra.ub.uni-muenchen.de/22295/. 20.11.2018.

[28] Hynninen S. M. (2007). Matching in Local Markets, Empirical Studies from Finland. Jyväskylä Studies in Business and Economics 56. University of Jyväskylä.

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

[30] Ibourk A., Maillard B., Perelman S. and Sneessens H. R. (2001). Aggregate Matching Efficiency: A Stochastic Production Frontier Approach, France 1990–1994, IZA Discussion Paper No. 339.

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

[32] Ilmakunnas P. and Pesola H. (2003). Regional Labor Market Matching Functions and Efficiency Analysis, *Labour, 17* (3*), 413−437*.

[33] Jeruzalski T. and Tyrowicz J. (2009). (In)Efficiency of Matching - The Case of a Post-transition Economy, MPRA Paper No. 16598, http://mpra.ub.uni-muenchen.de/16598/, 03.06.2019.

[34] Khitri W., Emrouznejad A., Boujelbène Y., Nejib M., Ourtani A. (2011). Framework for performance evaluation of employment offices: a case of Tunisia. *International Journal of Applied Decision Sciences 4, 16−33*.

[35] Khodabakhshi M., Lotfi F. H. and Aryavash K. (2014). Review of Input Congestion Estimating Methods in DEA, Hindawi Publishing Corporation, *Journal of Applied Mathematics*, *Volume 2014*.

[36] Kneip A., Simar L. and Wilson P. W. (2008). Asymptotics and consistent Bootstraps for DEA Estimators in Nonparametric Frontier Models, *Econometric Theory 24*, *1663–1697*.

[37] Kneip A., Simar L. and Wilson P. W. (2015). When bias kills the variance: Central limit theorems for DEA and FDH efficiency scores, *Econometric Theory 31, 394−422*.

[38] Kneip A., Simar L. and Wilson P. W. (2016). Testing Hypotheses in Nonparametric Models of Production. *Journal of Business & Economic Statistics, 34: 3, 435-456*.

[39] Lehmann, H. (1995). Active Labor Market Policies in the OECD and in Selected Transition Economies, World Bank Policy Research Working Paper No. 1502, Washington D.C.

[40] Orme, C. and Smith, P. (1996). The potential for endogeneity bias in data envelopment analysis. *Journal of the Operational Research Society*, 73–83.

[41] OSF, Official Statistics of Finland: Employment Service Statistics [e-publication]. Helsinki: Ministry of Economic Affairs and Employment [referred: 14.6.2019]. Access method: http://www.stat.fi/til/tyonv/index_en.html.

[42] Petrongolo B. and Pissarides C. (2001). Looking into the Black Box: A Survey of the Matching Function. *Journal of Economic Literature XXXIX, 390–431*.

[43] Pissarides, C. A. (1990). Equilibrium unemployment theory. Oxford: Oxford University Press.

[44] Politis D. N., Romano J. P. and Wolf M. (2001). On the asymptotic theory of subsampling. *Statistica Sinica 11, 1105–1124*.

[45] Ramirez J., Vassiliev A. (2007). An efficiency comparison of regional employment offices operating under different exogenous conditions. *Swiss Journal of Economics and Statistics 143, 31–48*.

[46] R Development Core Team (2008). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. URL: http://www.R-project.org.

[47] Santin D. and Sicilia G. (2017). Dealing with endogeneity in data envelopment analysis applications. *Expert Systems With Applications,* 68, 173–184.

[48] Recruitment of labor (2010–19), by several authors. Ministry of Economic Affairs and Employment, Analyses. https://tem.fi/julkaisulista?subject=TEM-analyysit (In Finnish).

[49] Riksrevisionen (2012). Effektivitetsmätning som metod för att jämföra arbetsförmedlingskontor. Granskningsrapport 2012: 9. (In Swedish).

[50] Sheldon G. (2003). The efficiency of public employment services: a nonparametric matching function analysis for Switzerland. *Journal of Productivity Analysis 20, 49–70*.

[51] Shephard R. W. (1970). Theory of Cost and Production Functions, Princeton: Princeton University Press.

[52] Sickles, R. C. and Zelenyuk, V., (2019). Measurement of Productivity and Efficiency, New York: Cambridge University Press.

[53] Simar L. and Wilson P. W. (2001). Testing restrictions in nonparametric efficiency models. *Communications in Statistics –Simulation and Computation, 30: 1*.

[54] Simar, L. and Wilson P. W. (2007). Estimation and inference in two-stage, semi-parametric models of productive efficiency, *Journal of Econometrics 136, 31–64*.

[55] Simar L. and Wilson P. W. (2008). Statistical Inference in Nonparametric Frontier Models: Recent Developments and Perspectives. In Fried H, C. A. Knox Lovell, and Shelton S. Schmidt: The Measurement of Productive Efficiency and Productivity Change. Published to Oxford Scholarship Online: January 2008, DOI: 10.1093/acprof:oso/9780195183528.001.0001.

[56] Simar, L. and Wilson P. W. (2010). Inference by the m out of n bootstrap in nonparametric frontier models, *Journal of Productivity Analysis 36 (1), 33–53*.

[57] Simar, L. and Wilson P. W. (2011). Estimation and Inference in Nonparametric Frontier Models: Recent Developments and Perspectives, *Foundations and Trends in Econometrics* Vol. 5, Nos. 3–4 (2011) 183–337.

[58] Simar, L., and Wilson P. W. (2020). Hypothesis testing in nonparametric models of production using multiple sample splits, *Journal of Productivity Analysis 53, 287–303*. https://doi.org/10.1007/s11123-020-00574-w.

[59] Talonen Markku and Tuomaala Mika (1994). The Effectiveness of Employment Offices, Labour policy studies No. 79, Ministry of Labour Helsinki. (In Finnish).

[60] Talonen Markku (1998). Overall results of public employment offices, Labour policy studies No. 191, Ministry of Labour, Helsinki. (In Finnish).

[61] Talonen Markku (2020). Efficiency of Regional Labor Markets: Evidence from Finland, Finnish economic Papers no 2/2020. https://www.taloustieteellinenyhdistys.fi/finnish-economic-papers-2-2020/ 19.5.2021.

[62] Talonen Markku (2021). Efficiency of TE-offices: A Stochastic Frontier Analysis. Finnish Labour Review 1/2021, Ministry of Economic Affairs and Employment. (In Finnish).

[63] Vassiliev A., Ferro Luzzi G., Flückiger Y., Ramirez J. V. (2006). Unemployment and employment offices' efficiency: what can be done? *Socio-Economic Planning Sciences 40, 169–186.*