I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13399

(In)Efficiency of Employment Offices: A Study on Welfare Benefits Determination – Is There a Trade-off between Time Saving Case Management and Quality of Decisions?

Katharina Dyballa
Kornelius Kraft

JUNE 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13399

(In)Efficiency of Employment Offices: A Study on Welfare Benefits Determination – Is There a Trade-off between Time Saving Case Management and Quality of Decisions?

Katharina Dyballa
TU Dortmund University

Kornelius Kraft
TU Dortmund University, IZA, KU Leuven and ZEW

JUNE 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

(In)Efficiency of Employment Offices: A Study on Welfare Benefits Determination – Is There a Trade-off between Time Saving Case Management and Quality of Decisions?*

This study investigates the efficiency of the process of benefit determination for welfare recipients in Germany. A stochastic frontier analysis is used to compute (in)efficiency of Jobcenter (employment offices) in terms of average processing time used for determining benefit levels per case. Next, the quality of the process of welfare benefits determination is considered by analyzing the share of upheld opposition because of misapplication of the laws. No effect of the (in)efficiency term on quality is estimated such that the quality of decision is unrelated to the time input. Manning of the employment offices seems to be to a large extent determined by other factors than a fair allocation of scarce resources in relation to demand for them. However in the case of treatment of one particular group (newly registered unemployed) and one organizational measure (offices that arrange specific appointments) a trade-off is estimated. Moreover, better skilled employees need less time for servicing cases and produce less erroneous decision.

JEL Classification: J65, I38, H53, H55, C54

Keywords: employment offices, efficiency, processing time, quality of decisions

Corresponding author:
Kornelius Kraft
Faculty of Business and Economics
TU Dortmund University
Vogelpothsweg 87
44221 Dortmund
Germany

E-mail: kornelius.kraft@tu-dortmund.de

* We thank Manuel Eder, Peter Hempel, Reno Richter and Rolf Stellermann for many very helpful discussions on the issues investigated by project. We also thank the Institute for Employment Reseach (IAB), Nuremberg for collection the data and BearingPoint for data preparation. Furthermore, we are very grateful to a referee and the editor for many very helpful and interesting comments.

1. Introduction

German employment agencies have two tasks: Job placement and basic income support for job-seekers[1]. The target to integrate employable persons into the labor market as fast and efficiently as possible has been the subject of many studies[2]. Aside of reintegration, in order to efficiently support unemployed people on their way back into labor, a welfare state's responsibility is also to provide an appropriate financial support for job seekers. This second task of employment agencies is largely neglected in empirical research, although this activity comprises the larger part of personal resources of the agencies (in Germany)[3]. This paper reports the results of an empirical study on the efficiency of the welfare determination process in terms of labor input per case and secondly in terms of the quality of the decisions.

In Germany, unemployed persons receive for 12 to 24 months (length depends on age) a fixed percentage (60-67% in dependence of private circumstances) of their last income and this payment is called unemployment benefits I[4]. The employment agencies ("Arbeitsagenturen") are responsible for this part of the unemployed. After that period basic social security and social assistance (officially called unemployment benefits II and in common parlance named Hartz IV after the chairman Peter Hartz of a commission, which proposed this reform of welfare payments) replaces unemployment benefits[5]. Whereas in other countries[6] the determination process of welfare benefits and the job placement of unemployed is mostly organized in different independent organizations, in Germany employment offices (called "jobcenters") are at the same time responsible for unemployment benefits II and employment services.

The employment offices usually separate their staff for the tasks of job placement and service provision. According to Bundestagsdrucksache 18/8956 (German Parliament Document No. 18/8956), a total of 45948 people were employed by the job centres in 2013. Of these, 23561

[1] The relevant legislative basis is the Zweite Buch des Sozialgesetzbuch (SGB II, Book II of the Social Code). It also defines the two relevant tasks of the employment offices.

[2] Examples of studies on efficiency of employment agencies with respect to placement of unemployed persons are e.g. for the US Cavin and Stafford (1985), for Switzerland Ramirez and Vassiliev (2007) Vassiliev et al. (2005) or for Sweden Andersson et al. (2014), Althin et al. (2010), Althin and Behrenz (2005, 2004).

[3] See Federal Employment Agency (2015).

[4] All regulations concerning social assistance for the unemployed and dependent relatives by unemployment benefits I are enacted in Book III of the Social Code (Sozialgesetzbuch III).

[5] All regulations concerning social assistance for the unemployed and dependent relatives by unemployment II are enacted in the already mentioned Book II of the Social Code (Sozialgesetzbuch II).

[6] In the U.S. the United States Employment Service (USES) is responsible for the provision of labor exchange and job finding assistance to job seekers and employers, whereas a joint state-federal program- so called "unemployment Insurances"- are responsible for the calculation and payments of the unemployed's cash benefits. Similarly Switzerland separates responsibilities such that employment offices (Regionale Arbeitsvermittlungszentren (RAV) are in charge of the job placement whereas an "unemployment fund" is responsible for the determination of unemployment benefits and how they are paid.

(51.28%) were responsible for job placement and 22387 (48.72%) for the provision of benefit payments. Only 912 employees were responsible for both tasks at a ratio of 50% to 50% (which are included in the above figures in equal parts). Since the administrative costs are financed jointly by the federal government and the municipalities in fixed proportions, it is not trivial to determine the total amount. The BIAJ estimates the total amount for 2013 at 5,264 billion euros[7]. Hence about 2.565 billion Euros administrative costs are caused by the process of determination of benefits for the unemployed and their families.

In 2013 about 33.68 billion Euros were spent for assistance (Federal Employment Agency 2014), which amounts to 1.19% of GDP (expenditures for active labor market policy not included). In that year the total number of persons supported by financial aid through German employment offices adds up to 6,126,322 persons[8]. Hence the topic of our study has some economic relevance.

The efficiency of the public service is the topic of several empirical studies and evidence of inefficiency in comparison to private suppliers of the same or similar services is a frequent result[9]. The reforms of the Schröder chancellorship among other items also aimed at improving the efficiency of employment offices. Therefore the German ministry for employment and social affairs (BMAS), the Federal employment agency, their local representatives, the Federal ministries and municipalities as well as employment offices fix target agreements like for example the establishment of efficient employment offices and compliance with legal norms (Matiaske et al., 2015, 146). However, in practice, almost nothing is known about the efficiency of jobcenters with respect to fulfilling this task and this is of particular interest as offices' employees complain about quite stressful work conditions[10]. Only 250 of the overall 96300 employees (in 2015) of the Federal Employment Agency receive financial incentives in term of bonus payments if certain goals are reached

[7] Bremer Institut für Arbeitsmarktforschung und Jugendberufshilfe (Bremen Institute for Employment Research and Youth Employment Assistance BIAJ) materials january 31,2020.

[8] In Germany, the serviced recipients are called "need communities (Bedarfsgemeinschaften)". In 2013 3,323,823 need communities received social assistance. A need community (common household) encloses individual recipients as well as dependent family members or permanent partners. The majority (2,419,804) of the need communities consists of just one recipient.

[9] Boyne (2003) summarizes 65 empirical studies on determinants of public service performance. The author analyses five potential sources of service improvements: resources, regulation, markets, organization and management. The most consistent influences on performance in public service have "resources" and "management".

[10] The so-called Bund-Länder Ausschuss (committee of the Federal Republic and Federal States), which is responsible for the governance and supervision of basic social security provision in Germany, initiated a research project aiming at evaluating the present situation at employment offices. Based on this evaluation, recommendations concerning the range of time necessary for determination of welfare benefits per case should be developed and the inefficient employment offices have to take measures to improve their present situation. This paper reports major parts of the scientific results of this study.

(Kaltenborn et al. 2010, 34). Hence in the absence of incentives for the overwhelming majority of the employees, inefficiency would not be surprising.

In detail, the study has the intention to identify the influence of economic variables on processing time for determining benefit level. We use a number of exogenous variables affecting time needed for servicing welfare recipients like for example the amount of new registrations, the amount of increased requires of the long-term unemployed or the qualification of the employees of the employment offices. By applying stochastic frontier analysis (SFA) we are then able to identify the determinants of the welfare determination process and the relative efficiency of employment offices simultaneously In particular, efficiency is estimated as the deviation from an efficient frontier taking account of the specific cost driving factors an employment office is confronted with.

The process of benefit determination is highly complicated, characterized by numerous legal requirements and under permanent revisions (see below). As a result many wrongful decisions are made. Consequently we identify an employment office's quality by the amount of inappropriate decisions in the welfare determination process. In particular we refer to the number of upheld oppositions (due to incorrect application of the law) divided by the total number of cases dealt by the employment office as our quality indicator. Finally, in order to analyze weather there is a trade-off between quality and efficiency we evaluate the impact of the estimated inefficiency term from SFA on the quality of decisions.

Previewing the results, unsurprisingly we find efficiency differences, but despite of missing incentives (or disadvantages in cases of insufficient performance) they seem not to be very large. The average employment office is 7.0% away from the cost minimum. Interestingly, the (in)efficiency term has no impact on the quality of decisions. Robustness tests with other econometric models support this conclusion. Furthermore, only for one variable (employment office arranges specific appointments) significant coefficients with different signs are found in the estimates for time use and upheld oppositions. Hence, for one organizational variable a trade-off between time input and the ratio of upheld oppositions seem to exist. However, qualification of the employees has both an effect on processing time needed for servicing and the ratio of inappropriate decisions. Thus, we conclude that better educated personnel would improve efficiency with respect to time use and quality of the decisions. With the aid of rough calculations based on the average pay of middle and upper grade employees, the cost increases of employment of members of the better qualified group can be calculated. This cost increase is compared with the savings in hours and upheld oppositions to the decisions on

benefit determination. However, a clear statement is not possible with the information available.

2. Literature Review

In general there is high political interest in evaluating quality and efficiency of public employment services. The existing literature usually evaluates efficiency by investigating the matching process of the unemployed with vacancies[11].

One of the first studies evaluating technical efficiency on employment office level was conducted by Cavin and Stafford in 1985. The authors examine the efficiency of 51 American State Employment Security Agencies in providing three different output categories: quality (average placement wage) targeting (the amount of successfully placed young applicants) and quantity (the amount of successfully placed adults). The authors refer to the efficiency term as cost efficiency by using an approach based on frontier production and cost functions[12]. As a result Cavin and Stafford (1958) provide evidence for quite large efficiency differences between employment offices ranging from 38 percent higher costs than expected in New York to 27 percent lower costs than expected in Florida.

The later studies can be grouped according to the applied method that is used in order to quantify efficiency: Thus, the first strand of literature uses non-parametric and non-stochastic frontier methods like Data Envelope Analysis (DEA) which are typically used in operation research disciplines (linear programming). The second strand of literature refers to stochastic production frontier analysis which, contrary to the first method, also controls for stochastic components while estimating efficiency.

One of the first studies applying DEA is Sheldon (2003). The author investigates efficiency in Switzerland as a term of a regional placement office's "matching" efficiency (speed with which jobless people become employed and at the same time vacancies attract job seekers.) By using Swiss data from 1997-98 on 126 regional placement offices he reveals that at the average placement offices only reach two third of their efficiency potential. A big part of inefficiency is due to the failure to exploit increasing returns-to-scale, meaning that the size of the placement offices should be bigger in order to handle the stocks of unemployed and job

[11] A recent study of Andersson et al. (2014) provides an overview on previous international studies on this topic. It turns out that the majority of existing literature refers either to Swedish or Swiss labor markets.

[12] The authors use corrected OLS (COLS) estimation technique in order to calculate an employment offices' deviation from the best performing office (best practice approach).

vacancies efficiently. Furthermore Sheldon (2003) points out that increased counselling of the unemployed is more effective in improving efficiency than any other labor market insturments. Vassiliev et al. (2006) also use DEA to analyze the Swiss labor market using data on 156 regional employment offices from 1998 to 1999. They focus on efficiency in terms of output maximization – in this case number of hires - from a given set of inputs. Similarly to Sheldon (2003), Vassiliev et al. (2006) show that given the average amount of inputs used in in Swiss employment offices it would be possible to increase the output by 15 percent. Further it turns out that the used office- and region- specific variables are able to explain one third of variation in an office's efficiency score indicating that the external operating environment significantly influences efficiency.

Althin and Behrenz (2005, 2004) [13] provide two studies on the efficiency and productivity of Swedish employment offices. They use output variables like open market jobs and job placements. By applying non-stochastic production frontier analysis (DEA) they focus on an input-minimizing model[14]. As a result an employment office's efficiency is measured as the mix and use of inputs in relation to the produced outputs (offices using less input in order to produce same or more output are more efficient). Efficiency varies a lot between Swedish offices and the mean efficiency measure has a value of just a bit more than .7, implying that output could be produced by nearly 30 percent lower use of input. Additionally, Althin and Behrenz (2004) used Tobit estimation in order to find an explanation for the variation in efficiency between the offices. It turns out that more unemployed and more vacancies have a significant positive effect on the efficiency scores whereas an office's municipalities population has no impact on the efficiency score. In the most recent study of Althin et al. (2010) the authors extend their model by considering intertemporal aspects of public employment services[15] while computing efficiency and simultaneously modelling and controlling for an office's expected work load[16]. Again the authors detect large differences in efficiency between the employment offices. Another application of DEA is Andersson et al. (2014), which is the most recent study on employment offices in Sweden [17]. They focus on an output based approach, having the target to maximize the output[18] for a given stock of input,

[13] Althin and Behrenz (2005) use data on 253 Swedish employment offices for the 1992-1995 periods whereas Althin and Behrenz (2004) use data on 297 Swedish employment offices in 1993.
[14] The authors refer to Farrell (1957).
[15] Intertemporal/intermediate outputs are reallocations over time and defined in three categories: unemployed gets a job that does not fulfill his wishes, unemployed is placed in training, client is openly unemployed.
[16] The expected work load is estimated with the help of duration models. It represents the time that is required in order to transform an unemployment registration to a final output.
[17] Instead of the traditional distance function the authors use a directional distance function approach.
[18] In this study the main output variable is represented by the number of individuals that got a job placement resp. the number of individuals that are transferred to outside education.

whereas the authors also consider intermediate outputs as future inputs (similar to Althin et al. (2010)). This study considers yearly data on 185 employment offices during the period from 2004 until 2010. Although there is evidence for general inefficiencies (between 7 and 10 percent for the observation period[19]), it turns out that they are absolutely seen smaller than in previous studies using Swedish data.

The second strand of the literature uses variants of stochastic or deterministic frontier analysis. Ramirez and Vassiliev (2007) is -from a technical perspective- the closest one to our approach. In comparison to the Vassiliev et al. (2006) the more recent version is extending the model (and also recent general literature) by using a parametric stochastic approach, the so called stochastic production frontier model. The authors model a classical production function and a production function based on the former mentioned matching function[20] (result do not differ much). By using Swiss monthly panel data from November 2000 until December 2001 there is significant evidence for the existence of technical inefficiency. Due to the application of a stochastic analysis the author state that 19.1%[21] of the efficiency variation is due to random noise, underlining once again the importance of not simply ignoring this aspect by using non-stochastic approaches. At the average employment offices generate 84% of the potential output given the input factors. The authors detect (in contrast to former literature) negative returns to scale implying that bigger offices are not necessarily more productive than smaller[22]. Based on German data Fahr and Sunde (2006) analyze the efficiency of the matching[23] process by using variation across 117 German labor market regions during the years from 1980 until 1997[24]. Evaluating efficiency with means of stochastic frontier analysis reveals that search effectivity is depending on the age and skills of the unemployed, i. e. the matching process is more efficient in regions with young labor market participants and in regions with higher fractions of high and low educated[25].

[19] Meaning that at the average the employment offices could increase the number job placement by 7 to 10 percent using the same amount of inputs.

[20] In both cases production is measured by using a proxy for hires as single output variable. The difference between the classical production function and the matching production function is that the later controls for open vacancies as an input into production frontier.

[21] Respectively 22.3% for the model considering a matching function.

[22] This is similar to our results. See below

[23] As mentioned before also in this context the efficiency of "matching" refers to the efficiency in matching unemployed with firms that aim to fill open vacancies.

[24] Note that this study is based on regional data. For Germany so far no study is available using employment office level data.

[25] Several studies investigate by use of German data the effects of public sector sponsored training on the successful reintegration of jobless people into the German labor market (i.e. Lechner et al. 2011, Fitzenberger and Völter 2007).

In principle the discussed studies measure on the one hand efficiency of the JobCenter (employment offices) and on the other hand the quality of the match between the unemployed and the job requirements[26]. In regions where a strong industrial restructuring takes place, reemployment will be rather difficult irrespective of the efforts of the employees of the employment offices. As expressed before, our interest is on the efficiency of the process of determining welfare payments. We estimate the determinants of processing time for servicing the welfare recipients taking account of specific cost driving factors. Next we are able to investigate the quality of the process by investigating the relation between time input and the ratio of upheld oppositions to decisions. Hence the main contribution of our study is firstly the direct estimation of the efficiency of the employment offices and secondly a test on the quality of this process.

One advantage of our evaluation method is the possibility of getting information on managerial efficiency within employment offices, by taking into account the characteristics of the persons being cared for and those of the labor market offices as well as regional labor market conditions. As stated above, German employment offices are (in contrast to other countries) simultaneously responsible for the job placement and the determination process of welfare benefits and an analysis of the efficiency of the determination process enables a direct test on managerial efficiency within German employment offices.

3. Research Questions and Hypotheses

In 2009 the Federal ministry for employment and social affairs made a recommendation on the relation of 130 cases (of welfare recipients) per employee of German employment offices responsible for servicing. The actual relation was at that time somewhat lower (1:115), and as the responsible ministries for employment and social affairs did not agree to this proposal, no effort was made to enforce the recommended service ratio. It is therefore not surprising that the employment offices realize quite different manning ratios.

On April 24, 2013, the federal and state committee decided to launch a research project to explain the main factors influencing the employment offices' use of personnel in determining

[26] There is literature on regional labor markets efficiency which does not consider an employment office's specific characteristics. This efficiency is mostly referred to as matching efficiency analyzing the amount of hires in a certain region (output of production) explained by open vacancies (input) in a certain region. For an overview of literature on the regional matching efficiency see Fitzenberger and Furdas (2012).

benefit levels for the unemployed. Furthermore, the achieved quality of service provision should be taken into account when determining the necessary personnel expenses[27].

As it will be shown in the next section, much of the variation in processing time used per case in the determination process of basic income support for job-seekers is caused by exogenous factors. Hence efficient allocation rules for personal might actually take account of differing requirements for servicing intensity. Whether the resources of the employment offices are really fixed in dependence of the number of cases and the specific requirements for intensity of consulting is an open question.

If resources are allocated efficiently according to the number of cases and their specific time requirements, the observed differences in efficiency between employment offices should be rather small. One reason for differences in time required per case could be managerial inefficiency. An alternative explanation for possible considerable and unexplained differences in time required per case could be simply arbitrarily determined resource allocations (coincidence, historical reasons, political decisions) without a regular evaluation process. We include and test the effects of variables related to additional working time like the newly registered unemployed or recipients with increased requirements and several organizational measures like quality management or accounting measures, which explicitly aim at improving efficiency. The determination of welfare payments is highly complicated and under permanent revision. In the time period between the introduction of the SGB II reform in 2005 and 2012 62 laws have been modified with in part drastic changes (Job Center NRW 2013). On average a file documenting the servicing of one household comprises 650 pages (Bundesvereinigung der Deutschen Arbeitgeberverbände 2015). Therefore, a certain share of incorrect decisions is not surprising.

The recipients have the possibility to oppose decisions if they suppose inaccurate decisions. In the first place oppositions are reviewed internally by a legal redress office. Inappropriate decisions are identified by the number of oppositions, which were approved by the redress offices. These are in particular oppositions where decisions were based on misapplications of laws. Thus, our variable to identify (low) quality of decisions is the number of upheld oppositions (due to incorrect application of the law) divided by the total number of cases dealt by the employment office.

A priori it is unclear, how efficiency in determination of benefit levels affects the quality of decisions. Clearly, one possibility is that time saving processing of the cases implies

[27] Beschlussempfehlung der AG Personal für den Bund-Länder-Ausschuss am 24. April 2013 (Decision recommendation of the Working Group on Personnel for the Federal-Government and Länder Committee on April, 24 2013)

inadequate decisions[28]. However, alternatively efficiency in one dimension of servicing the recipients (determination of payouts) might be positively correlated with efficiency in terms of accuracy of the decisions. For example organizational measures might affect both dimensions of the process of determining welfare payments. Finally, of course efficiency in benefit determination might be uncorrelated with the quality of the decisions. If this were true, the use of disproportional much time for the determination process would have no benefit in any respect and consequently manpower of overstaffed employment offices could be shifted to other JobCenter or used for totally different purposes.

4. Data

The data for this study is on 299 Job Center (employment offices) operated by the Federal Employment Agency and the municipalities as *Joint Local Offices* ("gemeinsame Enrichtungen") and covers the year 2013[29]. These offices were formed in 2010 after a decision of the Federal Constitutional Court in Germany in 2007, which decided that the former institution (ARbeitsGEmeinschaften, short ARGE) was not in accordance with the German constitution[30]. After a change of the legal foundation in 2010, including a modification of the German constitution, new institutions were launched. The municipalities (counties and larger cities) were granted the possibility to choose between forming Joint Local Offices or introducing Approved Local Providers ("kommunale Einrichtungen") conducted by the municipalities only. 104 approved local offices were accordingly created in addition to 304 joint local offices (of which 299 participated in this study).[31]

The administrative costs of the employment offices operating as joint local offices are shared. In particular the municipalities cover 16% while the Federal Employment Agency bears 84% of the expenses. Expenditures for the recipients are shared as well. The Federal Employment Agency bears the expenses on the so-called normal requirements ("Regelbedarf"), covering basic needs of living. Expenditures for rents and heating are taken over by the municipalities. This evaluation is based on the joint local offices and not the approved local providers as the latter ones were not subject to the supervision and direction of the Bund-Länder-Ausschuss.

[28] Concerning the second task of employment offices Hofman et al. (2012) found out that a better caseworker-to-client ratio in counselling and job placement activities in German employment offices leads to more job placements.

[29] Unfortunately only data from one year is available and therefore no panel models can be appplied.

[30] See Konle-Seidl (2009) on the reasons for this decision.

[31] For a comparison of the performance between Approved Local Providers and Joint Local Providers with respect to transition probabilities of the unemployed into employment see Holzner and Munz (2013).

Information concerning the characteristics of employment offices includes processing time used for servicing and the number of upheld oppositions because of decisions due to misapplication of laws. On this basis our two dependent variables are formed. Exogenous variables are available on the characteristics of the employment offices, the welfare recipients and regional data. The variables are defined and discussed in detail below.

Since 2004 the employment offices are responsible for the financial support of unemployed persons as well as of persons which live together with them in a common household if the unemployed is covered by ALG II. These are spouses, partners and children who do not work and are not registered as unemployed themselves. Therefore the number of persons serviced is quite high. In contrast to the ALG I which is provided for the first 12 to 24 months of unemployment, the determination of ALG II is independent of former income.

Within the employment offices the tasks concerning determination of welfare benefits and employment services are separated. In this study only determination of financial support is analyzed.

The so called Bund Länder Ausschuss (Federal Government and Länder committee) decided in 2012 that a study on the benchmarks for the provision of services on benefit determination by the employment offices is necessary. This involved recording the working hours as precisely as possible as the essential component of the provision of services.

In the first place average processing time used by a Jobcenter (employment office) to deal with all cases is estimated. For this purpose a central task catalogue was developed in cooperation with representatives of the Jocentrs (employment offices). As a result of various workshops 42 tasks were identified, which were divided into 6 task groups.

All employees of the employment offices responsible for benefit determination were asked to report the hours used for the 6 tasks exposed above and 86% of all employees actually participated in the survey. For the remaining employees working time was estimated on the basis of what has been reported by employees with comparable tasks. Public and individual holidays as well as individual absences due to illness were taken into account. The responses were checked for plausibility, and in cases of (very few) missing data points imputations were implemented. The collected data was sent back to the employment offices for a final approval. On this basis averages for every task group and every employment office were computed. These data was subsequently used to calculate our dependent variable.

Table 1 summarizes how the employees total working time in hours can be divided according to the following six tasks and the respective percentage:

Table 1: Task groups for the employees responsible for benefit determination

	Percentage
• hours used for global tasks[32] ($h1$)	11.10
• hours used for handling of new applications ($h2$)	14.43
• hours used for permanent case management ($h3$)	55.44
• hours used for cross-case tasks[33] ($h4$)	11.25
• hours used for education and social participation ($h5$)	2.50
• hours used for tasks outside the specific range ($h6$)	5.28

The overview shows that permanent case management is the most important task. As already mentioned above, the tasks of determining benefits and employment placement are separate in Germany. This is also clear from the above list, as the specialist staff only spend a small part of their time (5.28%) on tasks outside the scope of determination of monetary support.
In order to derive precise information on the hours used for the determination of benefits per case ($hBenefitDet$) we have to modify the employees' total working time. Tasks two to four focus explicitly on the determination of benefit levels and are hence entirely part of $hBenefitDet$. As tasks five and six are not directly connected to the benefit determination process, we do not consider them for the calculation of $hBenefitDet$. Task one represents global tasks which are generally included. However, due to the fact that task five and six are not included in the calculation, we reduce $h1$ by a proportional amount that would refer to task five and six. Finally the average hours used for the determination ($hBenefitDet$) in ageny i is defined as follows:

$$hBenefitDet_i = \frac{h1_i - \left(\frac{h5_i + h6_i}{h2_i + h3_i + h4_i + h5_i + h6_i} * h1_i\right) + h2_i + h3_i + h4_i}{cases} \quad \textbf{(1)}$$

The denominator is the total number of cases. The literal denomination of the unit serviced is "Bedarfsgemeinschaft" (need community). However, in order to preserve clarity we use the word "case". In our sample the average number of persons served per case is 1.88. The

[32] Global tasks are defined as: management tasks, basic tasks, general tasks, personnel and organisational development, training and supervision of initial staff, professional supervision, complaints management.
[33] Cross-case tasks are defined as: data maintenance and correction lists, approval of benefit determination, mail processing and record keeping in the area of benefit determination.

variable $hBenefitDet_i$ is the dependent variable in the first part of this study. It is used after a logarithmic transformation.

A number of explanatory variables are used to explain $hBenefitDet$. The variables can be divided into measures characterizing the labor market, the serviced unemployed and their families, the employment offices and particular organizational measures applied to improve efficiency[34]. Table 2 presents an overview on all variables and the corresponding descriptive statistics.

The characteristics of the local labor markets are firstly included by the number of unemployed persons, divided by the total population below 65 (variable is called *Unemployment*). The hypothesis is that a higher share of unemployed leads to routine in dealing with the cases and lower costs per case. Secondly, the number of newly registered benefits recipients in 2013 divided by total recipients of benefits enters the regression (*Newly Registered*). Usually cases, which have to be serviced for the first time, create additional work.

The characteristics of the serviced unemployed are taken up by several variables. The structure of the benefits recipients is included by the number of long-term recipients (four years or more) divided by all recipients of benefits (*Long-term Unemployed*). On the one hand such cases may require less time for servicing, as perhaps the individual circumstances do not change much over time. In contrast, it is sometimes said that servicing the long-term unemployed is particularly time intensive as these persons are frequently unhappy with their situation and have specific needs like debt counselling, addiction counseling or psychosocial care. The ratio of cases with increased requirements (pregnant women, handicapped persons, single parents, persons who need specific diets because of health problems) to all cases will in all likelihood increase time costs (*Increased Requirements*). Similarly the variable persons per case (*Persons*) may lead to increased processing time, as every separate member of a need community has an entitlement for financial support. The unqualified unemployed perhaps need more time for assistance and therefore the ratio of the number of these persons divided by the total number of unemployed (*Unqualified*) is added as well[35].

[34] Please note, we use for the explanation of hBenefitDet and the ratio of upheld oppositions the same set of variables, although the relevance may differ. By this procedure we want to test whether a trade-off between time input and quality of decisions may be identified for specific variables (aside of the effect of the inefficiency term).

[35] In calculating this rate, only the number of long-term unemployed is taken into account, i.e. the number of unemployed people who are actually serviced by the Job Centers.

2: Variable Definition and Descriptive Statistics

Variable name	Notation	Mean	Std. Dev
Dependent Variables:			
hBenefitDet	hours used for the determination of benefits per case	16.35618	2.673355
Oppositions	number of oppositions, which were upheld after internal control because of a misapplication of the laws, in relation to the total number of cases dealt with	.0233411	.0133605
Labor market Characteristics:			
Unemployment	number of unemployed persons, divided by total population below 65 in a given region	.0768199	.0479718
Newly Registered	ratio of number of newly registered recipients to total number of recipients in 2013	.5239529	.1277693
Unemp19962013	Growth of unemployment between 1996 and 2013	-.2872391	0.1636731
Characteristics of the welfare recipients:			
Long-term Unemployed	ratio of cases of long-term recipients (four years or more) to all cases	.5660257	.0789751
Increased Requirements	ratio of cases with increased requirements like pregnancy, disability, single parents, health problems to all cases	.3282743	.0934839
Persons	number of persons per case	1.828027	.1056844
Foreigners	ratio of foreign recipients to total number of recipients	.1669691	.0931784
Income	ratio of recipients with income from other sources to total number of recipients	.64278	.0259707
Unqualified	ratio of unqualified unemployed to all unemployed (both measured for our relevant SGB II group only)	.3814373	.0960155
Characteristics of the Jobcenter (employment offices):			
Employees	number of full-time equivalent employees	80.20236	101.2312
Lower and Intermediate	ratio of employees on the lower and intermediate level to total number of employees	.6410166	.194146
Organizational characteristics:			
One Caseworker	dummy, takes unit value if cases are handled by just one person	.0952381	.294044
Scheduled	dummy, takes unit value if employment office arranges specific appointments	.4659864	.4996923
Accounting	dummy, takes unit value if specific personal is specifically responsible for management and operating accounting	.5068027	.5008062
Organization	dummy, takes unit value if personal is specifically responsible organizational design	.2857143	.4525242
Quality	dummy, takes unit value if personal is specifically responsible for the existence of quality assurance	.4013605	.4910095

It is possible that servicing of foreign unemployed person implies more time input as language knowledge and lack of familiarity with German institutions lead to increased time need. The ratio of foreign recipients to the total number of recipients is applied to take this

possibility into account (*Foreigners*). A significant part of the recipients have some income from other sources like e.g. income from (part-time) working. This will in all likelihood imply higher handling time and is included by the share of recipients with income to all recipients (*Income*).

Specific characteristics of the employment offices are considered by two variables. The logarithmic value of the number of employees (*Employees*) tests for economics of scale. In addition qualification differences might matter. Better qualified employees probably need less time for processing cases (and simultaneously produce less inappropriate decisions). The employees are classified into four levels, which in turn are based on the basic education level: lower service, intermediate-level, upper-intermediate-level and upper-level. The relative shares of employment of the four groups are shown in table 3.

Table 3: Ratios of Employees According to Education

level of education	ratio	sd
lower	0.010	(0.025)
intermediate	0.634	(0.192)
upper-intermediate	0.355	(0.196)
upper	0.001	(0.001)

Standard deviation in parentheses

As only very few employees of employment offices are assigned to the lowest and the highest levels of the hierarchy, empirically relevant are only the employees belonging to the intermediate and the upper-intermediate levels. The intermediate-level employees have a practical administrative training, while the upper-intermediate level employees have studied at a university of applied sciences (Fachhochschule) for at least three years. There exist specialized universities of applied sciences for the public service and the Federal Employment Agency itself runs “universities of applied labor studies”. The (few) employees assigned to the upper level have studied for at least four years at a university[36] and have earned a master degree or a state examination (Staatsexamen), which prepares for working in the public sector on the upper level.

The structure of the employees in employment offices is included by the share of employees on the lower and intermediate level (*Lower and Intermediate*). The hypothesis is that better educated employees are able to fulfill their tasks more efficiently and if this were true the inclusion of the share of employees on the lower or intermediate-level would lead to positive coefficients in both the working time and the upheld opposition equations.

[36] Universities of applied sciences are practice-oriented, while universities offer the traditional academic education.

A set of variables takes up organizational characteristics. One variable is a dummy, which has unit value if cases are handled by just one person (*One Caseworker*). Division of labor would lead to specialization of employees and given that the process is quite complicated this might improve speed of decisions and quality alike. To the contrary in bureaucratic organizations with imperfect coordination and no incentives a negative impact on productivity could also exist. Another organizational variable focuses on time scheduling. Some offices arrange specific appointments for the unemployed (dummy variable *Scheduled*), while in other offices the applicants simply have to wait until they get served.

Next, explicit initiatives to improve organizational efficiency are included. These are identified by dummy variables which assume unit value, if personal is specifically responsible for efficiency improving tasks. The efficiency measures considered are presence of management and operating accounting (*Accounting*), existence of organizational design (*Organization*) and existence of quality assurance (*Quality*). All information was collected and provided by the Institute for Employment Research (Nuremberg).

It is possible that, for historical reasons, employment offices may differ in the equipment with employees. This may be the case, for example, because in the past there was a particularly large or small number of unemployed people in the district in question, but the situation has since changed. However, the number of caseworkers does not necessarily have to have been adjusted accordingly. This hypothesis is taken into account with a variable that measures the growth of unemployment over the period 1996 to 2013 (*Unemp19962013*).

5 Stochastic frontier analysis on time used for determination of welfare support

The analysis of time input used to determine the level of welfare benefits can be regarded as an estimation of a cost function. One way to identify the relative efficiency of the employment offices at performing their tasks is by use of stochastic frontier analysis. Stochastic frontier analyses are parametric estimation models, which investigate the relation between input and output. Typically, production or cost functions are estimated[37]. The procedure estimates simultaneously coefficients for the explanatory variables as well as deviations from an efficient frontier[38]. We use a cost function, although our dependent variable does not actually represent a monetary value, but rather working time. However, working time is paid differently depending on the training. This is taken into account by a

[37] As mentioned earlier Ramirez and Vassiliev (2007) also use stochastic frontier analysis in order to evaluate employment offices' efficiency. In contrast to us they estimate a production function instead of a cost function.
[38] A detailed presentation of the methods and models can be found in Kumbhakar and Knox Lovell (2000).

variable that reflects the average level of training. The econometric model is in the case of a cost function:

$$\ln(hBenefitDet_i) = \beta_0 + \sum_{j=1}^{k} \beta_j \ln(x_{ji}) + v_i + u_i \qquad \textbf{(2)}$$

The dependent variable $\ln(hBenefitDet_i)$ is the logarithmic value of the time needed per case, $\ln(x_{ji})$ are logarithmic values of the k explanatory variables[39], β_0 represents the constant term and β_j stands for the k coefficients of the exogenous variables. The stochastic term of the equation is denoted by v_i and the inefficieny term by u_i. In accordance with most applications of stochastic frontier analysis we assume in the case of v_i a symmetric normal distribution. However, as the inefficiency term u_i is non-negative (but also normally distributed), the total residual $\varepsilon_i = v_i + u_i$ is asymmetric[40]. Given that the dependent variable is specified in logarithmic values, the term $\exp(u_i)$ stands for relative (in)efficiency. Larger values of $\exp(u_i)$ represent lower efficiency.

In contrast to the standard interpretation of the residuals as a measure for all unknown factors not considered in the estimation, here the inefficiency term is of central interest. Intuitively, the inefficiency effect must increase costs, and therefore only positive values of u_i are allowed. This creates the asymmetry, which is the major difference to OLS (and other common estimation methods). The stochastic frontier analysis models are usually estimated by Maximum Likelihood. Given that only an estimate of the error term ε_i is possible, the decomposition of v_i and u_i is realized by an approximation (based on the assumed distribution, like e.g. the normal distribution)[41]. The values of u_i can then be computed by use of the structural parameters estimated by the model.

[39] Stochastic frontier analysis is based on the use of logarithmic variables. The variable ShareLowMedium has in one case a value of zero. To realize logarithmization, instead of zero the lowest observed value is inserted, and this observation gets in addition a specific dummy variable.

[40] In the case of a production function, however, the total residual is $\varepsilon_i = v_i - u_i$, as output is usually below the maximum and in cost functions expenditures are usally above the efficient frontier.

[41] See Greene (2012, ch. 19.2.4) .

Table 4: Stochastic Frontier Analysis on Determinants of Processing Time (loghBenefitDet)

	(1) lhBenefitDet b/se	(2) lhBenefitDet b/se
Unemployment	-0.137***	-0.125***
	(0.030)	(0.031)
Newly Registered	0.230**	0.241**
	(0.099)	(0.107)
Long-term Unemployed	0.379**	0.340*
	(0.156)	(0.176)
Increased Requirements	0.100***	0.089***
	(0.027)	(0.027)
Persons	-0.285	-0.315*
	(0.175)	(0.182)
Foreigners	-0.005	-0.016
	(0.020)	(0.022)
Income	0.249	0.279
	(0.266)	(0.280)
Unqualified	0.019	0.037
	(0.043)	(0.048)
Employees	0.015	0.025
	(0.014)	(0.016)
Lower and Intermediate	0.059***	0.060***
	(0.014)	(0.014)
One Caseworker	0.014	0.032
	(0.026)	(0.025)
Scheduled	0.029**	0.032**
	(0.013)	(0.014)
Accounting	0.004	0.007
	(0.015)	(0.016)
Organization	-0.011	-0.010
	(0.016)	(0.018)
Quality	-0.006	-0.001
	(0.015)	(0.016)
Unemp19962013		-0.044
		(0.046)
_cons	3.07***	3.036***
	(0.25)	(0.267)
lnsig2v	-4.448***	-4.778***
	(0.243)	(0.126)
lnsig2u		-4.614***
	-4.980**	(0.679)
	(0.917)	
N	294	270

Notes: * p < 0.1, ** p < 0.05, *** p < 0.01, robust standard errors in parentheses.

Table 4 displays the stochastic frontier estimation results of two specifications. The results in column (1) are based on the efficiency measure e^{u_i} and the results in column (2) are estimated using the alternative measure u_i .

The unemployment ratio (*Unemployment*) on the regional level decreases time input per case[42]. Apparently a higher ratio of unemployed persons increases routine with servicing such cases. The ratio of long–term unemployed persons (*Long-term Unemployed*) has a positive coefficient in model one, implying that the long term unemployed need a more careful support. Unsurprisingly people which have particular requirements (*Increased Requirements)* need more time for servicing. Similarly handling of persons, who ask for the first time for support (*Newly Registered*), takes additional time. Surprisingly, the variable person has a negative coefficient in the specification (2) which is significant at the 10% level. This would express that more persons per case would lead to lower time needed for the benefit determination. Employees without higher education (*Lower and Intermediate*) need more time for handling the cases. Working on the basis of appointments (*Scheduled)* increases the time needed per case. Dealing of cases by just one caseworker (*One Caseworker*) has no impact. The measures to explicitly improve efficiency (*Accounting*, *Organization*, and *Quality*) have unfortunately no effect on processing time (but wait for the effect on quality). The ratios of recipients with income (*Income*) and of the unqualified unemployed (*Unqualified*) have no effects on time needed for servicing.

The size variable (*Employees*) has no impact on the dependent variable These results are in line with Vassiliev and Ramirez (2007), who even detected decreasing returns to scale of employment offices (in the context of placement of unemployed persons), suggesting that a bigger employment office does not necessarily lead to higher productivity.

The variable *Unemp19962013* is insignificant. This is also the case if the development of unemployment over the periods 1985-2013, 2000-2013 and 2005-2013 are used as variables.

One possible way to express the importance of the stochastic component versus the (in)efficiency is the relation $\lambda = \frac{\sigma_u}{\sigma_v}$. It is simply the ratio of the standard deviations of the two components. Interestingly this value differs considerably between specification one and specification two. The random component is somewhat larger ($\lambda = 0.81$) the inefficiency term. Table 5 summarizes the estimates on (in)efficiency.

[42] Stochastic frontier analysis does not compute R^2 statistics. To give an impression of the relevance of the exogenous variables, the regressions have been repeated by OLS. The R^2 is .then .53.

Table 5: Summary Statistics of (In)Efficiency Measures

	mean	sd	min	max
hBenefitDet	16.356	2.673	10.749	28.692
u_i	0.067	0.024	0.023	0.207
e^{u_i}	1.070	0.24	1.024	1.23

The minimal and maximal values point to a considerable range. This result is in line with previous literature which also detects considerable variation in efficiency between different employment offices (i. e. Cavin and Stafford (1985), Althin and Behrenz (2005, 2004)). The average deviation from the theoretical cost minimum is 7.0% and the difference to the employment office with lowest realized costs is about 4.6%. There is no obvious point of comparison for these figures, but if we take the empirical studies surveyed in the literature section, these deviations from the efficiency frontier seem to be of moderate magnitude. Computing monetary values leads however to values of 179 million Euros if the difference to the theoretical minimum is computed and 117 million Euros if the difference to the observed minimum is the reference point[43]. These amounts are not negligible.

6 Quality Assessment

The main purpose of the estimation of the stochastic frontier analysis model is the generation of a variable concerning efficient time allocation, which takes into account the particular burden an employment office has to deal with. Next, the computed (in)efficiency term is used to explain the employment offices' service quality.

To investigate precision of decisions we use data on the number of upheld oppositions of welfare recipients against decisions. An opposition is filed in 22.80% of all cases. Oppositions are in the first place examined internally by a legal redress office. We use the number of oppositions, which were upheld after this internal control because of a misapplication of the laws, in relation to the total number of cases dealt with[44]. This happens in 2.33% of all cases[45]. The number of granted appeals (called *upheld oppositions*) is our dependent variable

[43] Taking 48.72% of the total administrative costs of 5,264 million Euros and multiplying this figure by values of .046 and 0.070 leads to figures mentioned in the main text.
[44] The objections can also be upheld on other grounds, such as the submission of additional documents.
[45] This means that about 10% of all oppositions are approved because the legal basis was not applied correctly.

in the second part of the empirical study. If the objection is not successful, the benefit recipient has the option of taking legal action against the decision on the objection[46].

The empirical model is basically a two stage procedure where efficiency terms are estimated in the first place. These estimates are used in the second step. In two stage approaches of this kind the conventional standard errors are not valid. Therefore we use bootstrapping with cluster adjustment and 200 replications. Note, this bootstrapping procedure is based on the inclusion of both parts of the estimation procedure for every bootstrap sample.

The dependent variable is log-transformed ($logupheldOpposition$). As in three cases zero upheld oppositions are observed, we follow the suggestion of Cameron and Trivedi (2009, 532) how to proceed in such a situation[47]. Because of the small number of censored observations we use OLS, but Tobit leads to almost identical results. All variables used in the estimation of time input efficiency are included here as well. The same variables are used in both equations to allow comparison of effects and to avoid omitted variable bias. However, this procedure is not without problems and is not used in section 7 robustness test.

In the case of several variable the reason is the obvious relevance for both questions like that the more complicated cases (Increased Requirements, Persons, Income) not only lead to more processing time, but with some likelihood also imply to a higher incidence of complaints[48]. The servicing by appointment (*Scheduled*) will probably lead to fewer problems while the effect of the handling by one caseworker (*One Caseworker*) is unclear. It is quite likely that the qualification level of the employees (*Lower and Intermediate*) has here an impact as well because better qualified personal is expected to do a better job. As just expressed organizational variables may increase efficiency with respect to the number of upheld oppositions, but the contrary is also possible, as organizational innovations might also increase stress and by this also costs in terms of more upheld oppositions.

It is possible that foreigners have problems with the whole process and do not complain about decisions. In contrast to this hypothesis, they might have the impression to be unfairly treated and therefore to the opposite more often require reexamination of decisions. Thus, no obvious hypothesis can be made with respect to the variable *Foreign* and this is also the case for the variables *Newly Registered*, *Employees*, *Unemployment* and *Long-Term Unemployed*.

[46] Unfortunately, we have no information on the number of granted appeals at the level of individual employment offices.

[47] The transformation requires that zero values of our dependent variable are adjusted to a value which is smaller or equal to the smallest uncensored value. In our case, this minimum is slightly lower than the smallest observed value.

[48] Exclusion of irrelevant variables does not affect the results.

Table 6 presents the results with column one referring to the specification using e^{u_1} and column two referring to the specification using e^{u_2} as explanatory variable. The ratio of employees with lower or medium level education (*Lower and Intermediate*) increases the share of upheld oppositions indicating that better qualified employees have an advantage in analyzing the extensive and complicated legal basis for the decisions.

Working on a scheduled basis (*Scheduled*) reduces the share of upheld oppositions. Apparently this way to service needs more time, but is also more careful. In a similar vein, the newly registered unemployed (*Newly Registered*) are on the one hand dealt with more processing time and on the other hand this allocation of resources leads to less inappropriate decisions.

The need communities with more members increase the ratio of upheld oppositions (*Persons*), presumably as these cases are more complicated and therefore lead to inappropriate decisions. Similarly, a higher share of foreign recipients (*Foreigners*) cause a higher share of upheld oppositions and this suggests that these cases are dealt with insufficient care.

Two of the variables aiming at improving organizational efficiency do not meet expectations. The use of accounting methods (*Accounting*) increases oppositions without having an advantage with respect to time input[49]. The same conclusion is true with respect to the variable *Organization*. In contrast Quality assurance (*Quality*) works as expected by reducing upheld oppositions.

Adding the inefficiency term (neither e^{u_i} nor u_i) has no significant effect. In addition simpler (in)efficiency measures have been tried as alternatives to stochastic frontier analysis. Firstly, simply average processing time used for servicing (*hBenefitDet*) is included. Secondly, the difference between the observed value *hBenefitDet* and the expected value $\widehat{hBenefitDet}$ is computed. The expected value is calculated on the basis of the specific parameter values of the explanatory variables and the coefficients. This method is called deterministic frontier model. The implementation of these two variables does not lead to significant coefficients either. Hence on the basis of upheld opposition concerning the decisions of the employment offices, there exists no trade-off between processing time and quality of the service (based on the inclusion of several cost drivers in the first equation).

[49] In general the efficiency of accounting systems and target agreements of employment offices are quite controversially discussed (e.g. Kaltenborn et. al. (2010, ch. 4.3. Matiaske et. al. 2015).

Table 6: Determinants of Ratio of Upheld Oppostitions to Total Number of Cases

	(1) logOppositions	(1) logOppoitions
Unemployment	-0.066	-0.066
	(0.143)	(0.143)
Newly Registered	-0.150	-0.150
	(0..387)	(0..387)
Long-term Unemployment	0.752	0.753
	(0.771)	(0.772)
Increased Requirements	-0.104	-0.104
	(0.150)	(0.150)
Persons	2.028**	2.028**
	(1.040)	(1.040)
Foreigners	0.189**	0.189**
	(0.092)	(0.092)
Income	-1.369	-1.367
	(1.230)	(1.230)
Unqualified	-0.959***	-0.959***
	(0.205)	(0.205)
Employees	0.089	0.089
	(0.070)	(0.070)
Lower and Intermediate	0.271***	0.271***
	(0.077)	(0.077)
One Caseworker	-0.203	-0.203
	(0.130)	(0.130)
Scheduled	-0.171***	-0.171***
	(0.066)	(0.066)
Accounting	0.150**	0.150**
	(0.079)	(0.079)
Organization	0.131	0.131
	(0.081)	(0.081)
Quality	-0.149*	-0.149*
	(0.078)	(0.078)
e^{u_i}	1.308	
	(46282)	
u_i		1.495
		(47117)
_cons	-7.861	-6.560
	(46310)	(9.426)
N	294	294

Notes: * p < 0.1, ** p < 0.05, *** p < 0.01, bootstrapped standard errors in parentheses

What attracts attention is that the ratio of employment offices' employees with lower or medium-level medium level education has twice a significant impact. This share has an elasticity with respect to time needed for servicing of 0.059 to 0.060 and an elasticity with respect to erroneous decisions because of misapplication of the laws of 0.271.

Of course, those employed in upper-intermediate services earn more than those at intermediate level. The exact calculation of the difference in income is not easy. Employees at the Job Centers are paid according to the public service tariff (TVöD) and the tariff of the

Federal Employment Agency (TA-BA)[50]. For the TVöD, 4 pay groups with 6 development levels are relevant for each of the two employment groups, that is, 24 possible pay levels per group[51]. If we calculate the average of all 24 possible pay levels for both groups and the income differential, we obtain a value of 1.41 for the year 2013. According to this "back of the envelope" calculation, the better qualified would earn 41% more.

This difference in income cannot be justified by the approximately 6% increase in productivity in terms of working time per case. However, a 1% increase in the proportion of lower-skilled employees increases the number of inadequate decisions by 0.27%. Since there are likely to be other undiscovered decisions, the actual productivity difference is even higher. A high number of upheld oppositions probably implies a high rate of too low benefit decisions. Under this assumption, a lower rate of incorrect benefit decisions is expensive for the jobcenter, but good for the recipients of benefits and justice. Aside of the wrong decisions that are revised because of an opposition, presumably not all inappropriate decisions are brought to the attention of the legal address offices and the number of wrong decision will then be higher than it is recorded in the data. Furthermore, in all likelihood the recipients will not oppose too high benefit levels (at least not by purpose). We have no information on the share of benefits, which are inappropriately fixed at a too high level. Hence, the problems related with the employment of suboptimal trained personnel will probably be larger than what our data tells us. Clearly, a monetary assessment is difficult.

The employment offices or the Bund-Länder Ausschuss (committee of the Federal Republic and Federal States) should examine whether employment of (higher paid) upper-intermediate-level civil servants (substituting the intermediate-level civil servants) are worth the connected additional expenditures for them in comparison to the extra costs caused by the higher processing time and the increased probability of wrongful decisions in the case of the less well qualified personnel.

7 Robustness Tests

The previously used system of recursive equations with identical variables is only valid if the residuals of these two equations are independent of each other. Otherwise, different models must be used. Two alternative estimation models are applied in the following: IV and SUR. For both estimation models we exclude some variables from the respective equations.

[50] The pay scale of the BA is not very transparent, so we concentrate on the TVöD.

[51] See for an explanation concerning the pay groups Bundesverwaltungsamt (Central Service Agency oft he Federal Government) 2018.

Furthermore, the equation for the explanation of working time per case (lhBenefitDet) no longer uses stochastic frontier analysis, which is based on the assumption of an asymmetric distribution of the dependent variable.

Table 7: IV Estimation Results on Processing Time and the Ratio of Upheld Oppostitions

	(1) lhBenefitDet	(2) logOppositions
logOppositions	-0.02	
	(0.044)	
lhBenefitDet		-0.401
		(-0.56)
Unemployment	-0.140***	
	(0.031)	
Newly Registered	-0.229**	
	(0.101)	
Long-term Unemployment	0.408***	
	(0.125)	
Increased Requirements	0.100***	
	(0.030)	
Persons	-0.259*	1.485
	(0.152)	(0.949)
Foreigners		0.120*
		(0.061)
Income	0.240	-1.095
	(0.254)	(1.197)
Unqualified		-0.950***
		(0.184)
Employees	0.017	0.119**
	(0.014)	(0.060)
Lower and Intermediate	0.065***	0.310***
	(0.017)	(0.081)
One Caseworker	0.011	-0.205
	(0.026)	(0.125)
Scheduled	0.024	-0.148**
	(0.015)	(0.066)
Accounting	0.007	0.163**
	(0.017)	(0.075)
Organization	-0.008	0.125
	(0.018)	(0.078)
Quality	-0.011	-0.167**
	(0.017)	(0.075)
_cons	3.038***	-5.164**
	(0.276)	(2.364)
First Stage F-Value	11.424	16.515
Overidentification Test χ^2	0.002(p=0.97)	1.992 (p=0.574)
N	294	294

Notes: Robust standard errors in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01standard errors in parentheses

We'll start with an IV model. As is well known, one or two variables are instrumented to avoid possible correlations between the original variables and the error term. In our case, there could be a simultaneous influence of the two dependent variables, so that the recursive model would not be adequate. For example, a careful review of cases could lead to a longer processing time combined with a lower number of upheld oppositions. This would lead to a feedback to the inefficiency term in the equation to explain upheld oppositions. Excluding the variable upheld oppositions from the equation for the explanation of the working time per case would also not be advisable, as an omitted variable bias could result. If the described relationship (or other reasons for feedback processes) is actually present, the coefficients would be inconsistent.

The variables *Foreigner* and *Unqualified* are excluded from the equation determining lhBenefitDet. The variables *Unemployment*, *Newly Registered*, *Long-term Unemployment* and *Increased Requirements* are excluded from the equation determining Opposition. Hence, when estimating the IV models, the variables *Foreigner* and *Unqualified* are used as instruments for oppositions, while the variables *Unemployment*, *Newly Registered*, *Long-term Unemployment* and *Increased Requirements* are used as instruments for lhBenefitDet.

The IV specifications are subject to a twofold review of their acceptability. On the one hand, the relevance of the excluded instruments is checked. For this purpose an F-value is calculated, which should exceed a critical value of 10 (Staiger and Stock 1997 rule of thumb) in order to reject the hypothesis that the instruments are not (sufficiently) correlated with the endogenous variable (weak instrument problem). On the other hand, a Sargan overidentifying restrictions test is performed to check whether the instruments are really exogenous. The null hypothesis is that the instruments are uncorrelated with the error term. If this hypothesis is not rejected, the instruments are valid and correctly excluded from the second stage estimation equation. The corresponding chi-square values and their probabilities are calculated and shown at the bottom of the tables. All results can be found in Table 7.

The IV estimates hardly change the results and conclusions. We do not find any correlation between the two dependent variables. The IV specification tests support the approach used here.

Next the two equations are estimated by seemingly unrelated regressions (SUR)[52]. This procedure is used when several equations are estimated to explain different variables. Each

[52] See e.g. Wooldridge (2010, ch.7)

equation could in principle be independently estimated with OLS. SUR is now based on the possible correlation of the residuals with each other, but there is assumed to be no correlation between the residuals and the explanatory variables (endogeneity). This explains the term "seemingly unrelated", since the only relationship between the equations is via the residuals. SUR then increases the efficiency of the estimation, provided that not in both equations exactly the same variables are used. We use the same specification as for the IV model.

Table 8: SUR Estimation Results on Processing Time and the Ratio of Upheld Oppostitions

	(1) lhBenefitDet	(2) logOppositions
Unemployment	-0.138***	
	(0.025)	
Newly Registered	-0.237***	
	(0.077)	
Long-term Unemployment	0.389***	
	(0.120)	
Increased Requirements	0.109***	
	(0.028)	
Persons	-0.267	1.724**
	(0.177)	(0.840)
Foreigners		0.120*
		(0.067)
Income	0.246	-1.438
	(0.233)	(1.05)
Unqualified		-0.980***
		(0.191)
Employees	0.015	0.147
	(0.014)	(0.047)
Lower and Intermediate	0.060***	0.287***
	(0.016)	(0.075)
One Caseworker	0.014	-0.213**
	(0.023)	(0.108)
Scheduled	0.027**	-0.157**
	(0.013)	(0.063)
Accounting	0.004	0.161**
	(0.016)	(0.074)
Organization	-0.012	0.123
	(0.019)	(0.089)
Quality	-0.007	-0.166*
	(0.018)	(0.085)
_cons	3.141***	-6.720***
	(0.179)	0.846)
Breusch-Pagan test	1.861 (p=0.173)	
N	294	294

Notes: standard errors in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01standard errors in parentheses

A crucial test in this connection concerns the significance of the correlation between the residuals. If significance were given, then there were indeed common unobserved factors. If no significance is found, both variables are determined independently of each other. In this context, a popular test is the Lagrange multiplier statistic developed by Breusch and Pagan (1980).

The results of the SUR estimate are very similar to the previous ones. The correlation coefficient for the residuals of the two estimation equations is insignificant. Again, with this method of estimation, we cannot find a relationship between the two outcome variables studied.

8 Conclusion

The topic of this paper is the analysis of efficiency with respect to processing time used to determine welfare benefits and the impact of estimated (in-)efficiency on the quality of this process. We find no empirical evidence for the initially plausible hypothesis that there is a trade-off between the two endogenous variables. The recursive model finds no effect of the inefficiency in the time taken to process the cases on the quality of the decisions. This result is confirmed by IV and SUR estimates. Similarly, for the explanatory variables with one exception, there are no coefficients with significant opposite signs in the two equations. The exception is the variable scheduled (arrangement of individual appointments for the unemployed), which increases the time taken to process the cases, but also reduces the incorrect application of the law.

Unfortunately, this study finds not much evidence that approaches to improve organizational efficiency are working[53]. With respect to processing time no effect is found. With regard to upheld oppositions, accounting and organization tend to increase the rate of cases with incorrect application of the law. In contrast quality management reduces this ratio. The better qualification of employees improves efficiency in both dimensions and consequences of these results should be investigated. Having said this, it is not possible to quantify the costs and revenues exactly with the available information. The difference in income between the two

[53] A recent study of Matiaske et al (2015) empirically focuses on the working conditions in employment offices. The authors find out that due to "New Public Management" and the higher amount of emotional work employment office employees experience higher levels of stress than other professional groups. This could be one potential reason explaining inefficiencies in employment offices and a potential starting point in order to eliminate inefficiencies.

qualification groups can only be calculated imprecisely and the monetary costs of upheld oppositions cannot be quantified at all.

Although millions of people depend on the fast and correct determination of welfare benefits and billions are spent for this service, this is to our knowledge the first empirical study on the efficiency of the whole process. The results show the relevance of this research topic as there seems to be room for improvement of labor allocation.

One way to introduce target-oriented resource allocation is by yardstick competition. E.g. in energy regulation the providers of network facilities are compared and evaluated during every regulatory period. Similarly to our analysis the basis of the evaluation is Stochastic Frontier Analysis (and Data Envelopment Analysis). By application of such a benchmark system and an appropriate incentive mechanism, it would also possible to increase efficiency over time for all employment offices.

References

Andersson, C., Mansson, J., Sund, K. (2014), "Technical efficiency of Swedish employment offices", in: Socio-Economic Planning Sciences, 48, p. 57-64.

Althin, R., Behrenz, L., Färe, R., Grosskopf, S., Mellander, E. (2010), „Swedish employment offices: A new model for evalutating effectiveness", in: European Journal of Operational Research, 207, p. 1535-1544.

Althin, R., Behrenz, L. (2005), "Efficiency and productivity of employment offices: evidence from Sweden", in: International Journal of Manpower, 26(2), p. 196-206.

Althin, R., Behrenz, L. (2004), "An Efficiency Analysis of Swedish Employment Offices", in: International Review of Applied Economics, 18(4), p. 471-482.

Bremer Institut für Arbeitsmarktforschung und Jugendberufshilfe (Bremen Institute for Employment Research and Youth Employment Assistance BIAJ) (2020), Materials January, 31.

Breusch, T., Pagan, A. (1980), The LM Test and Its Applications to Model Specification in Econometrics, Reviere of Economic Studies, vol. 47, 239-254.

Bundestagsdrucksache (German Parliament Document) (2018), Document No. 18/8956.

Bundesvereinigung der Deutschen Arbeitgeberverbände (Confederation of German Employers' Associations **(BDA) (2015),** "Die Grundsicherung effizienter und erfolgreicher gestalten", Position Paper on enhancement social security provision according to SGB II for unemployed persons.

Bundesverwaltungsamt (Central Service Agency of the Federal Government 2018), Allgemeine Tätigkeitsmerkmale für den Veraltungsdienst, Cologne.

Boyne, G. A. (2003), "Sources of Public Service Improvement: A Critical Review and Research Agenda", in: Journal of Public Administration Research and Theory, 13(3), p. 367-394.

Cameron, A. C. and P. K. Trivedi (2009), Microeconometrics using stata, fifth edition, College Station, TX: Stata press.

Cavin, E., S., Stafford, F. P. (1985), Efficient Provision of Employment Service Outputs: A Production Frontier Analysis", in: The Journal of Human Resources, 20 (4), p. 484-503.

Fahr, R., Sunde, U. (2006), "Regional dependencies in job creation: an efficiency analysis for Western Germany", in: Applied Economics, 38(10), p. 1193-1206.

Farell, M. J. (1957), "The measurement of productivity efficiency", in: Journal of the Royal Statistical Society, 120, p. 253-281.

Federal Employment Agency (Bundesagentur für Arbeit) (2014), "Arbeitsmarkt in Zahlen" (Labor Market in Figures), Nuremberg, Germany.

Federal Employment Agency (Bundesagentur für Arbeit) (2015), "Haushaltsplan" (Budget), Nuremberg, Germany

Federal Government and Länder Committee (2013), Beschlussempfehlung der AG Personal für den Bund-Länder-Ausschuss am 24. April 2013 (Decision recommendation of the Working Group on Personnel for the Federal-Government and Länder Committee on April, 24 2013).

Fitzenberger, B., Furdas, M. (2012), "Benchmarking Regions: Estimating the Counterfactual Distribution of Labor Market Outcomes", IZA discussion paper, 6465.

Fitzenberger, B., Völter, R. (2007), "Long-Run Effects of Training Programs for the unemployed in East Germany", in: Labour Economics, 14(4), p. 730-755.

Greene, W.H. (2012), Econometric Analysis 7Th edition, Boston: Pearson

Hendry, D.F. and Santos, C. (2006), "Regression Models with Data-Based Indicator Variables", in: Oxford Bulletin of Economics and Statistics, 67, p. 571-595.

Hofmann, B., Kupka, P., Krug, G., Kruppe, T., Osiander, C., Stephan, G., Stops, M., Wolff, J. (2014), "Beratung und Vermittlung von Arbeitslosen: Ein Literaturüberblick zu Ausgestaltung und Wirkung", in: Sozialer Fortschritt, 11, p. 276-285.

Hofmann, B., Krug, G., Sowa, F. Theuer, S., Wolf, K. (2010), "Modellprojekt in den Arbeitsagenturen, Kürzere Arbeitslosigkeit durch mehr Vermittler, IAB-Kurzbericht.

Hofmann, B., Krug, G., Sowa, F. Theuer, S., Wolf, K. (2010), Wirkung und Wirkmechanismen zusätzlicher Vermittlungskräfte auf die Arbeitslosigkeitsdauer – Analysen auf Basis eines Modellprojektes, Zeitschrift für Evaluation, vol. 11 6-38.

Holzner, C., Munz, S. (2013), "Should local public employment services be merged with local social benefit administration?", in: Journal of Labor Market Research, 46, p. 83-102.

Jondrow, J.C., Lovell, C., Materov, I., Schmidt, P. (1982), On the estimation of technical efficiency in the stochastic production function model, Journl of Econometrics, vol. 19, 233-238.

Kaltenborn, B., Wielage, N., Henkel, A.I., Weihaupt, T., Buchholz, H.-L., Oschmainsky, F., Bothmer, A. V. (2010), "Zielsteuerung in der Arbeitsverwaltung – ein europäischer

Vergleich", Research report No. 409 for the German Ministry for Employment and Social Affairs (BMAS),

Konle-Seidl, R. (2009) Neuregelung der Jobcenter für Hartz-IV Empfänger (New regulations of employment offices for welfare recipients), Wirtschaftsdienst, vol. 89, 813-820.

Kumbhakar, S. C. und Knox Lovell, C.A. (2000), Stochastic Frontier Analysis, Cambridge: Cambridge University Press.

LAG NRW (2012), "Die Arbeitssituation in den Leistungsbereichen der Jobcentern in NRW,

Lechner, M., Miquel, R., Wunsch, C. (2011), „Long-Run Effects of Public Sector Sponsored Training in West Germany", in: Journal of the European Association, 9(4), p. 742-784.

Matiaske, W., Olejniczak, M., Salmon, D., Schult, M. (2015), "Arbeitsbedingungen in Jobcentern nach dem SGBII (gemeinsame Einrichtungen) – Mitarbeiterbefragung zu Arbeitsumfeld und psychischer Belastung", in: Industrielle Beziehungen, 22(2), p. 142-166.

Ramirez, J. V., Vassiliev, A. (2007), "An Efficiency Comparison of Regional Employment Offices Operating under Different Exogenous Conditions", in: Schweizer Zeitschrift für Volkswirtschaft und Statistik, 143(1), p. 31-48.

Sheldon, G. M. (2003), "The Efficiency of Public Employment Services: A Nonparametric Matching Function Analysis for Switzerland", in: Journal of Productivity Analysis, 20, p. 49-70.

Staiger, D. and Stock, J.H. (1997), "Instrumental Variables Regression with Weak Instruments", in: Econometrica, 65, p. 557-586.

Vassiliev, A., Luzzi, G. F., Flückiger, Y, Ramirez, J. V. (2006), "Unemployment and employment offices' efficiency: What can be done?", in: Socio-Economic Planning Sciences, 40, p. 169-186.

Wooldridge, J. M. (2010), Econometric Analysis of Cross Section and Panel Data, second edition, Cambridge, Mass.: MIT press.