Linnæus University
Sweden

Economics, degree project

Technical efficiency of Swedish district courts

- *a stochastic distance function analysis*

Author: Ajsuna Ragipi Rushid
Supervisor: Jonas Månsson
Examiner: Håkan Locking
Term: VT18
Subject: Economics
Level: Masters
Course code: 4NA06E

Linnæus University
Sweden

Abstract

The aim of this study is to measure Swedish district courts' technical efficiency for the period between 2000 and 2016 by applying the stochastic distance function approach. Although a very important issue from a policy perspective, a few studies have measured the efficiency of the courts. The narrow literature is also limited to using nonparametric methods, such as the DEA. The stochastic distance function has not been used for this purpose before and hence this is the first study to do so. The estimated mean score of technical efficiency is 93%. However, this study observes that efficiency levels increase throughout the studied period. Large variations between efficiency levels of different courts are also observed. Policy recommendations are to learn from courts with higher efficiency levels.

Keywords

Technical efficiency, Stochastic distance function, Courts efficiency

Acknowledgments

I would like to give my thanks to my supervisor Jonas Månsson. I am very grateful for his guidance and valuable comments on this thesis.

Linnæus University
Sweden

Table of contents

1 Introduction

It is widely known that public sector efficiency is fundamental for the smooth functioning of the state. The efficiency of the judiciary is at the heart of public sector efficiency as criminal justice is pivotal for a society. Even more fundamental is the efficiency of the courts, as they represent the spine of the judicial system and has been proven to affect the economy as a whole (Jappelli et al. 2005). However, balancing efficiency and quality in the judiciary has been a challenge for many countries. For instance, many EU countries still struggle to improve the efficiency of their courts, such as Spain and Italy (CEPEJ, 2016). The effectiveness of the Swedish courts has been no exception, with few reports indicating that there is room for performance improvements (Statskontoret, 2007; SNAO, 2017).

Although a very important issue from a policy perspective, little research has been done concerning the efficiency of the judicial system. Lewin et al. (1982) canonical study was the first to estimate judicial efficiency. The study used the non-parametric Data Envelopment Analysis (DEA), as has the vast majority of the studies estimating courts efficiency after their study. However, DEA has the disadvantage that it does not allow for hypothesis testing unless a two-step procedure is followed (Simar & Wilson, 2007). This raises an obvious need for a wider methodological variation within the field of judicial efficiency. For instance, the stochastic parametric approach allows for statistical testing of hypothesis and the construction of confidence intervals and could produce more reliable results.

The main objective of this study is to estimate Swedish district courts' technical efficiency for the years 2000-2016. The Swedish judiciary has experienced a number of reforms the last decades. Through a government decision in 1999, the district courts underwent a substantial organisational transformation, with the aim to improve efficiency. The reform implied merging of smaller district courts into larger ones, mainly due to the belief that larger courts have greater judicial and administrative power and could operate more efficiently (Statskontoret, 2007). The merging process resulted in a closedown of 48 district courts since 1999, a decrease of 50% in the number of district courts.

The long span of the yearly data will show whether the efficiency of the courts has improved or declined over time. The first half of the period is characterised by the merging process as all merges happened between 2000 and 2009. The second half period represents the post-reform period as no further merges happened between 2010 and 2016. Moreover, the analysis will be performed using three different outputs reflecting three different types of resolved cases and four inputs reflecting labour and capital used in the production of justice. Consequently, since a multioutput product is available, the analysis will be conducted using the stochastic distance function (Shephard, 1970). Technical efficiency will, hence, be represented by a non-negative error term that corresponds to the output distance function and represents the distance from the output to the production frontier. The proposed stochastic distance function approach has been previously used to study efficiency in a number of sectors, for instance in the public sector such as hospitals (Ferrari, 2006; Hamidi 2016 and for Swedish hospitals Löthgren, 2000) and railways (Coelli & Perelman, 2000), in the banking sector (Das & Kumbhakar, 2012) and agriculture (Irz & Thirtule, 2004). To the best of my knowledge, no previous study has employed the stochastic distance function approach when measuring the efficiency of courts. Hence, this study will be the first to employ the stochastic distance function approach and make a significant contribution to the efficiency literature.

The research documented in this paper has, therefore, a twofold purpose. Firstly, the aim is to estimate the Swedish district courts' technical efficiency using the stochastic distance function approach. Secondly, it aims to observe how efficiency levels develop over time and thereby determine whether the decision to merge smaller courts into larger ones has produced increases or decreases in technical efficiency.

The paper is structured as follows. It starts with an introduction to the Swedish judicial system and then with a background of the research concerning courts efficiency. Section 4 introduces the productivity and efficiency theory, as well as the distance function. It then discusses the data used as well as the chosen inputs and outputs. The stochastic output distance function approach is discussed in 5, and how it's used to estimate courts technical efficiency. The results are presented in section 6 and the findings are discussed in section 7. Section 8 concludes with policy recommendations.

2 The institutional setting – The Swedish judicial system

The aim of the Swedish judicial system, as with any judicial system, is to ensure the rule of law and legal security for Swedish citizens. The judicial organization consists of 80 different authorities and boards, with the courts being the core of the organization (Ministry of Justice, 2015). The Ministry of Justice holds the responsibility for the judicial authorities as well as for legalisation in the areas of civil law, constitutional law, procedural law and penal law (Ministry of Justice, 2015). The joint term for the Swedish court system is "Swedish courts" which consist of the general courts, the general administrative courts, the regional rent and tenancy tribunals, the Legal Aid Authority and the Swedish National Courts Administration (Ministry of Justice, 2015).

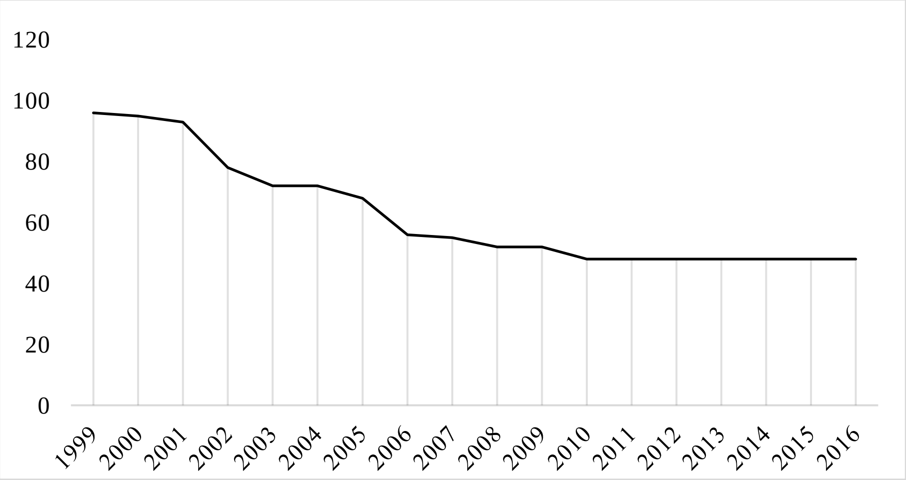
In general, there are three types of courts. The general courts consist of the district courts, courts of appeal and the Supreme court. The administrative courts consist of the administrative courts, administrative courts of appeal and the Supreme Administrative Court and the special courts which include for instance the Migration Court, Labour Court etc. Further, the instrument of Government specifies that courts are independent and that the Government or any other governmental agency cannot intervene in an individual case (Ministry of Justice, 2015).

This study will examine the technical efficiency of the district courts. The district courts are the first instance courts in Sweden unless otherwise stated. There are 48 district courts (as of 2018) that consider criminal, civil cases and other kinds of matters. Criminal cases are cases that are punishable. If there is any suspicion of a crime, after a police investigation, the prosecutors decide whether to bring the matter to court and prosecute. The civil cases deal with either property or family disputes (Statskontoret, 2007).

Moreover, the courts have a local connection, meaning that the cases they receive come from citizens in those municipalities that fall under the district court's jurisdiction (Swedish Courts, 2017). There are also land and environment courts at five district courts (the district courts of Nacka, Umeå, Vänerborg, Växjö, and Östersund) which handle cases concerning environmental, property, planning and building issues (Ministry of Justice, 2015). The district courts' legally trained judges should consist of a Chief judge and one or more judges. The court staff consists of Junior judges and Law clerks as well as court secretaries and officials with administrative tasks (Statskontoret, 2007).

The size of courts can vary, from around 300 employees in Stockholm to only 10 in Lycksele (SNAO, 2017). Due to a government action in 2000, the district courts underwent a substantial reorganisation. Smaller courts were merged into larger units, mainly due to the belief that larger courts have greater judicial and administrative power and could operate more efficiently. Another reason for the mergers is that courts' specialisation would increase (Statskontoret, 2007). Figure 1 shows the development of the merging process. To date, the reconstruction resulted in 48 district courts, a decrease of 50% from 1999. The biggest changes happened during 2000-2003 with 15 courts closing only in 2001. The last wave of mergers happened in 2009 when five district courts became one (SNAO, 2017). Since 2010, no further mergers between district courts occurred.

Figure 1. Number of district courts in Sweden for the period 2000-2016



Source: Own calculations, data from the Swedish National Courts Administration (SNCA).

3 Literature review

There is large and growing literature examining public sector productivity and efficiency. However, only a few studies have focused on the judicial system and its performance.

The earliest studies estimating courts efficiency were conducted in the 1990s (Lewin et al. 1982; Førsund & Kittelsen, 1992; Tulkens 1993 and Pedraja-Chaparro & Salinas-Jimenez, 1996). Lewin et al. (1982) were the first to study the efficiency of criminal courts and judicial districts and it is only since their work that judicial efficiency studies started to emerge. The study is conducted in North Carolina with the number of judgements and number of cases pending less than 90 days as outputs and the caseload[1], number of districts attorneys and stuff, days of a court held, number of misdemeanours within the caseload and the size of with population as inputs. The analysis is conducted using the Data Envelopment Analysis (DEA) approach and finds that 11 out of 30 district courts are inefficient. Kittelsen and Førsund (1992) explore Norwegian district courts' efficiency using yearly data for the period 1983-1986. The study examines the efficiency of 107 district courts with seven outputs grouped according to seven crime categories and two inputs, the number of judges and other remaining staff. The authors observe an average efficiency of 90 per cent and blame mainly the sizes of the courts to be the explanation for the efficiency levels.

[1] The caseload usually consists of pending and new cases.

The most common methodological approach used to assess judicial efficiency is the DEA. Tulkens (1993) is a rare example of a judicial productivity study not using the DEA method. More specifically, Tulkens (1993) uses the nonparametric Free Disposal Hull (FDI) method to evaluate courts productivity. The study is conducted using Belgian data for 187 courts for the period 1983-1985. The analysis shows that courts efficiency has room for improvements.

Another study measuring courts' efficiency is Chaparro & Jimnez (1996). Using Spanish data, the authors employ the DEA method with the staff and the judges as inputs. Due to lack of data, the authors distinguish the outputs based on whether the cases ended with a judgement or with other forms of resolution, such as settlement, withdrawal of suit and so on (Chaparro & Jimnez, 1996). However, this aggregation could be criticised for not taking into account the heterogeneity between different cases. Nevertheless, their findings suggest a substantial scope for improvements in courts performance, as only 23 per cent of the studied courts showed to be technically efficient (Chaparro & Jimnez, 1996).

Most of the scientific research has been addressed to individual national systems, although there are few comparative international studies as well (see Deyneli, 2012). The performance of the Italian judicial system has been explored several times, as its poor performance has been highlighted repeatedly by international organizations, for instance by CEPEJ (2008) and the World Bank (2009). Marselli & Vannini (2004) were the first to examine the efficiency of the Italian district courts using a two-stage DEA method. This approach first evaluates the decision-making units' (DMUs)[2] efficiency and later regresses the DEA efficiency scores on appropriate covariates (Marselli & Vannini, 2004). They use the number of resolved cases as output and the total number of cases and the judges' workload as inputs. Likewise, Castro & Guccio (2014) use the same two-stage DEA method, only with two outputs (resolved cases and other cases) and three inputs (judges, staff and pending cases). In the second stage, they add four covariates in order to investigate factors that influence technical efficiency. An average of 25 per cent inefficiency is found among the 27 Italian courts which is according to the authors mainly due to demand factors. In a later study (Castro & Guccio, 2018) the same authors assess the potential efficiency gains after the merging of Italian courts into larger ones. The study uses the same DEA method with four inputs: judges, administration staff, civil caseload and criminal caseload; and two outputs: resolved civil cases and resolved criminal cases. The results indicate that the Italian district courts are, on average, largely technically inefficient (Castro & Guccio, 2018).

[2] DMUs could be firms that produce tangible goods and services that are sold in the market or in the non-market sector or public bodies and the national economic sector (Bezat, 2011). In our case, a DMU is a district court.

Finally, Falavigna et al. (2015) employ a directional distance function with the DEA approach when estimating the technical efficiency of the Italian judiciary. The court delay is seen here as an undesirable output since courts solve disputes while simultaneously producing unavoidable delays (Falavigna et al., 2015).

Schneider (2005) carry out a study on German courts, also using the two-stage DEA analysis. The study examines how the German labour court system shapes courts performance. More specifically, Schneider (2005) looks at how the judges' education and careers affect courts productivity. In the first step, a DEA analysis is conducted with the number of cases as the output and number of judges and their caseloads as inputs. In the second step, Schneider (2005) investigates whether the share of judges that hold a Ph.D. will increase courts efficiency. Using pooled yearly data[3] for nine Courts of Appeal, the author finds that judges education does indeed increase courts efficiency. In a similar analysis, Deyneli (2012) examines the relationship between judicial efficiency and the salary of judges in 22 European countries. The study uses a two-stage DEA analysis, with the number of resolved cases as output and number of judges and staff as inputs. In the second stage, the judges' salaries, their training and the number of courts are used as determinants of courts inefficiency. The author finds a positive and significant relationship between judicial efficiency and salaries of judges (Deyneli, 2012).

Furthermore, to determine Brazilian courts efficiency, Yeung & Azevedo (2011) applies the DEA technique to 27 state courts. Two outputs are used, the number of first and second instance judgments, and two inputs, the number of judges and staff. The study finds an astonishing average efficiency of only 42 per cent (Yeung & Azevedo, 2011).

Santos & Amando (2014) have analysed the Portuguese judicial system by using the DEA technique and focusing on the performance of 223 first instance courts during the period of 2007-2011. The outputs used in this study are grouped based on case type, which makes 43 outputs in total. The inputs used are the number of judges and other staff. The authors discover that less than 16 per cent of the 223 studied courts are efficient, with the rest of the courts being not technically efficient (Santos & Amando, 2014).

A study focusing on the US, or more specifically, Florida's judicial efficiency is Ferrandino (2014). The study is using data on 20 criminal district courts with the number of judges as input and number of resolved cases per judge as output. The study employs the DEA method and finds that only a few of the courts operate efficiently.

[3] Schneider (2005) tries to solve the problem of using data on only nine courts by pooling the data and hence increasing the number of observations.

A more recent study by Major (2015) applies the DEA technique to measure the efficiency of the Polish courts. The analysis is carried out for the 26 district courts of the Cracow ward. The output is measured in number of resolved cases and the input in number of judges, number of assistants and number of officials. The DEA analysis shows that the capacity of the courts allows them to shorten the queue of pending cases (Major, 2015).

In their research of Swedish district courts, Mattson et al. (2018) refine the use of DEA by introducing the bootstrap option, as suggested by Simar & Wilson (2007). Using similar data as this current study, the authors employ DEA analysis to calculate the Malmquist productivity index of 48 district courts in the period of 2012-2015. The study uses number of judges, number of clerks and other personnel as labour inputs and the total amount of office space as capital input. The outputs are aggregated into resolved civil cases, resolved criminal, real estate and environmental cases and resolved matters[4]. Their findings indicate a 1.7 per cent decline in total factor productivity measured as a yearly geometric mean. Additionally, a substantial variation between courts is found, with 36 to 57 per cent of the courts having a negative change in total factor productivity while 16 to 36 per cent of courts having a positive change (Matsson et al. 2018). Their policy conclusion is that there is a considerable scope for improvements in the Swedish judiciary and that less efficient district courts can learn from more efficient ones.

It is evident that much of the available literature deals with the choice of inputs and outputs. For instance, Schneider (2005) argues that the inclusion of the caseload i.e. the justice demand, is important for the analysis since judges cannot provide their services without lawsuits being filed. Hence, omitting the caseload would imply that productivity is underestimated for those years in which courts are charged with a small caseload (Schneider, 2005). This view is supported by the abovementioned Kittelsen & Førsund (1992) and by Nissi & Rapposeli (2010).

Another common occurrence in the literature is to include number of employees (Lewin et al., 1982) as an input variable that measures labour or to distinguish between number of judges and other remaining staff (Kittelsen & Førsund, 1992; Chaparro & Jimnez, 1996; Schneider, 2005; Azevedo & Yeung, 2011; Deyneli, 2012; Ferrandino, 2014; Castro & Guccio, 2014; Santos & Amando, 2014). For the output side, it is common to separate the cases into type, for e.g. criminal cases and civil cases (Kittelsen & Førsund, 1992; Tulkens, 1993; Deyneli, 2012; Santos & Amando, 2014; Mattson et al. 2017 and Castro & Guccio, 2018).

To conclude, it is now well established that the vast majority of the studies use DEA when measuring courts efficiency. Overall, these studies highlight the need for more

[4] This study will use the same distinction between inputs and outputs, see section 4 for a more in detail discussion of the choice of variables.

variation in the choice of method that will enable an objective assessment of courts' functioning. The parametric stochastic distance function has been previously used to study efficiency in many areas, for instance in the public sector such as hospitals (Ferrari, 2006; Hamidi 2016 and for Swedish hospitals Löthgren, 2000), railways (Coelli & Perelman, 2000), in the banking sector (Das & Kumbhakar, 2012) and agriculture (Irz & Thirtule, 2004). However, the applicability of this method has not yet been proven on courts data. As previously mentioned, the present study will be the first to study courts efficiency using this method. Coelli & Perelman (1996) argue that the stochastic distance function can accommodate the multiple output and multiple input technology. Additionally, the method does not require price information thereby making it even more adequate for an analysis of district courts. The stochastic distance function, as well as the concept of technical efficiency, will be explained in more detail in the chapter that follows.

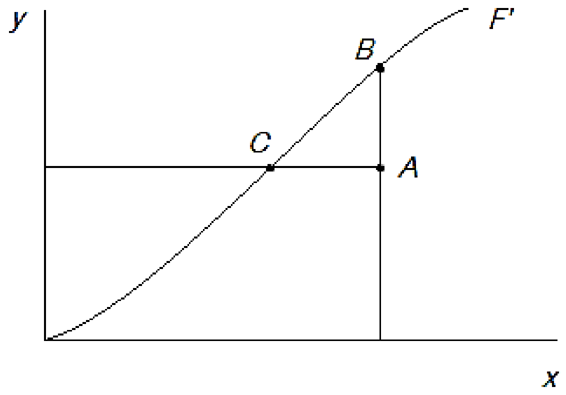
4 Theory

4.1 Technical Efficiency

The terms productivity and efficiency are often used interchangeably, despite the fact that the two terms represent two distinct concepts. In economic theory, the first concept is defined as the ratio of output(s) that are produced by a firm and the input(s) that it uses in the production. When one refers to productivity, one refers to the total factor productivity i.e. the productivity measure involving all factors of production such as labour, capital etc (Coelli et al. 2005). Efficiency, on the other hand, can be defined in two ways: allocative efficiency and technical efficiency (Farrell, 1957). A unit, in this case, a court, is allocative efficient if inputs are used in correct proportions, given their prices and marginal products (Coelli et al. 2005).

The concept of technical efficiency is easier to understand when one considers the production process in which inputs (*x*) are used to produce outputs (*y*) (Coelli et al. 2005). The curve *F*' in Figure 2 represents a production frontier that describes the relationship between the output and the input. Courts operate technically efficient if they are on that frontier (points *B* and *C*) or technically inefficiently if they are underneath the frontier (point *A*). Courts in point *A* operate inefficiently because technically they could increase output to point *B* without having to increase input (Coelli et al. 2005). Further, the area between the production frontier *OF'* and the x-axis represents the feasible production set, meaning that it contains all the input-output combinations feasible to courts given the production technology (Kumbhakar et al. 2015). The measure of technical efficiency can be either output- or input-oriented. A production plan is technically inefficient (output-oriented) when a higher level of output is technically achievable for the given inputs. Due to courts nature, this study will focus on the output-oriented measure. Hence, the objective is to not to reduce input usage but to produce an optimal amount of outputs given the inputs. This will be further discussed in the Method chapter.

Figure 2. Production Frontiers and Technical Efficiency



Source: Färe & Primont (1995)

4.2 The distance production function

When applying econometric approaches to estimate efficiency, it is necessary to specify an adequate functional form. Distance functions are suitable when we have a production function accommodating a multiple input and multiple output technology and we want to measure efficiency. The concept of a distance function was first introduced by Shephard (1953) and the rationale of the function is closely related to production frontiers (Coelli et al. 2005). In the following, the distance function will be discussed in more detail.

Distance functions can describe a multi-input, multi-output production technology without specifying behavioural assumptions such as how much of each input goes into the production of a certain output or requiring price data (Coelli et al. 2005). A distance function can be specified with either input or output specification (Kumbhakar et al. 2015). As mentioned in the previous section, this study will focus on the output distance function due to district courts nature.

A production technology transforming inputs into outputs can be represented by a technology set. This technology set will list the technologically feasible combinations of inputs and outputs. This can be shown by defining the firm's production technology[5]. Consider a vector of M inputs

$$x = (x_1, x_2, \dots, x_M) \quad (1)$$

and a vector of N outputs

$$y = (y_1, y_2, \dots, y_N). \quad (2)$$

The technology set can then be defined by

$$T = \{(x, y) : x \in \Re_+^N, y \in \Re_+^M, x \text{ can produce } y\}. \quad (3)$$

The technology satisfies the axioms listed in Shephard (1970) and Färe (1988). Further, for each input vector *x*, consider a technology production $P(x)$ defined as a set of feasible output vector *y*, that is obtainable from the input vector *x*:

$$P(x) = \{y: (x, y) \in T\}. \quad (4)$$

For multiple outputs and multiple inputs, the output distance function can algebraically be defined as

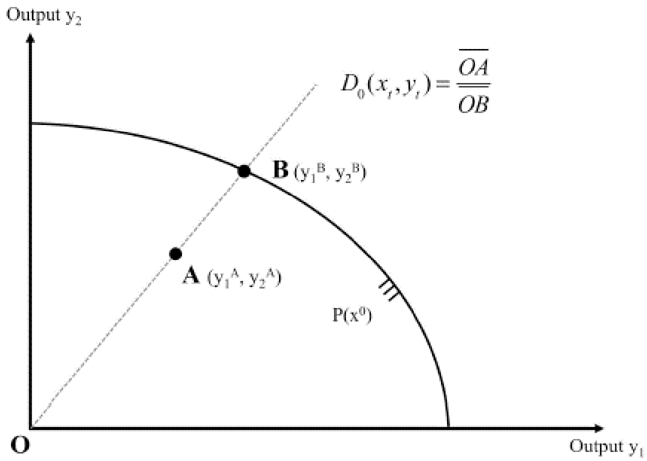
$$D_o(x, y) = \min.\left\{\Psi > 0: \frac{y}{\Psi} \in P(x)\right\} \quad (5)$$

where $P(x)$ describes the sets of outputs vectors that are feasible for each input vector *x* (Färe & Primont, 1995). The output distance function $D_o(x, y)$ is nondecreasing, positively linearly homogenous and convex in *y*, and decreasing in *x* (Lovell et al. 1994). Figure 3 illustrates the distance function in output space for the case of two outputs, y_1 and y_2. The boundary of the output set is equivalent to the production possibility curve. The distance function gives the relation of a given output vector (in Figure 3, $A = (\hat{y}_1, \hat{y}_2)$) to the maximal possible output with unchanged output mix (Färe & Primont, 1995).

Hence the function is defined as the maximum feasible expansion of the output vector with the input vector held constant. That is, given an input vector *x*, the value of the output distance function $D_o(x, y)$ places $y/D_o(x, y)$ on the outer boundary of $P(x)$ (Färe & Primont, 1995). Furthermore, the distance function, $D_o(x, y)$ will take a value which is less than or equal to 1 if the output vector *y* is an element of $P(x)$ i.e. is an element of the feasible production set.

[5] The following section is based on calculations and formulas presented in Färe & Primont (1995).

Figure 3. Output distance function for two outputs



Source: Färe & Primont (1995)

The distance function will, however, take a value of unity if y is located on the outer boundary of the production possibility set (Färe & Primont, 1995).

Since both the input and output distance functions are functions of x and y, they can be separated by the homogeneity restrictions (Kumbhakar et al. 2015). For instance, if we impose linear homogeneity restrictions on y, we can rewrite $D = f(x, y)$ as

$$\frac{D}{y_1} = f\left(x, \frac{y_2}{y_1}, \dots, \frac{y_n}{y_1}\right). \qquad (6)$$

The resulting function becomes an output distance function. Hence, homogeneity helps to separate the two distance functions.

4.3 The Translog Output Distance function

The translog function is a flexible functional form which allows to model complex features of the production function (Kumbhakar et al. 2015). Once one imposes the linear homogeneity conditions as in *(6)*, the translog output distance function can be derived (Kumbhakar et al. 2015). This is done in Kumbhakar et al. (2015) with first rewriting the distance function $D = f(x, y)$ as

$$D_o y_1^{-1} = f(x, \tilde{y}) \text{ where } \quad \tilde{y} = \left(\frac{y_2}{y_1}, \dots, \frac{y_m}{y_1}\right) \qquad (7)$$

and then taking log on both sides in order to obtain $\ln D_o - y_1 = \ln f(x, \tilde{y})$. The translog form of $f(x, \tilde{y})$ is then

$$\ln D_o - \ln y_1 = \beta_0 + \sum_m \beta_m \ln x_m + \sum_n \gamma_n \ln \tilde{y}_n + \frac{1}{2}\left[\sum_m \sum_k \beta_{mk} \ln x_m \ln x_k + \sum_n \sum_l \gamma_{nl} \ln \tilde{y}_n \ln \tilde{y}_l\right] + \sum_m \sum_n \delta_{mn} \ln x_m \ln \tilde{y}_n \tag{8}$$

The translog output distance function can be transformed into a stochastic one by simply adding a two-sided noise term v. In this way, we get an estimable equation in which the error term is $v + u$ (Kumbhakar et al. 2015). This is furthermore discussed in chapter 6. The next chapter describes the data used in this investigation.

5 Data

The data for the existing analysis is obtained from the Swedish National Courts Administration (SNCA). The SNCA offers detailed statistics on total 102 district courts with the number of courts decreasing throughout the sample due to mergers[6]. The data are yearly data for the period 2000-2016. This will allow to observe courts technical efficiency during the merging process. The main advantage of using panel data is that we consider some heterogeneity that may exist by introducing a time-invariant, court-specific effect (Kumbhakar, 2015 p. 241). Another advantage of using panel data is that we can examine whether inefficiency is persistent over time or time-varying. However, a limitation with our dataset is that some courts exist only for a short period of time due to the merging process. This implies that efficiency levels are estimated only for few years for merged courts, whereas efficiency for non-merged courts is estimated for every year of the studied period. Hence, the dataset is an unbalanced panel encompassing 102 district courts from 2000 to 2016.

The dataset has previously been studied in Mattsson et al. (2018) [7]. To determine district courts efficiency, Mattsson et al. (2018) aggregated the outputs into three groups and decided four input variables based on previous research and economic theory but also on what the courts have stated to be reasonable resources and performance measures in interviews conducted by the Swedish NAO (SNAO, 2017).

[6] For instance, when two courts merge, one court disappears from the sample.
[7] However, Mattsson et al. (2018) estimate courts efficiency only for the years 2012-2015 whereas this study's data span between 2000 and 2016.

This study will follow Mattsson et al. (2018) choice of input variables. In the following, a more detailed discussion of the input and output variables is presented.

5.1 Inputs

The choice of inputs must reflect the labour and the capital used in district courts production of justice. Since courts are in general labour-intensive, labour constitutes a major component of the total expenditure on inputs in courts[8]. The most common measure of labour input is the total number of employees at one court. Since the data allows us to distinguish between different types of employees, the labour input can be divided into three groups: number of judges, number of law clerks and other personnel. All labour input variables are measured as full-time equivalent workers.

Judges (DOM) are the employee category that resolves many of the district court's cases. This group includes permanently working judges but also temporarily working judges such as student judges that serve during their education, judges working when the caseload is larger than usual and retired judges that work when needs arise.

The second group of employees is the law clerks (NOT). This group includes the law clerks and the drafting law clerks. The law clerks work as judges' trainees and usually determine some types of less-complicated cases. The drafting law clerks are usually qualified lawyers that prepare cases and make proposals for decisions.

The third labour input variable is the remaining personnel (AD) employed by the district courts. This input variable includes employees with administrative tasks such as court officials and secretaries but also security guards and cleaners.

To measure capital input, this study will use the office space of each court as 13% of the total expenditure of district courts went to rental expenses in 2016 (Swedish Courts, 2016). These capital costs are additionally important to include in our analysis since rental expenses have probably decreased with time due to mergers. If this prediction is true, then we would find a further increase in technical efficiency as fewer inputs are used in district courts production of resolved cases. Hence the fourth input variable is the office space measured in area (AREA).

5.2 Outputs

Courts main task is to produce justice. Therefore, it seems straightforward that the number of resolved cases is the output measure of courts. However, aggregating different types of cases with different resource requirements into one single measure may produce aggregation errors. Hence, it is inappropriate to aggregate criminal cases

[8] 71% of the total expenditure of Swedish district courts went to labour expenses in 2016 (SNCA, 2016).

with civil cases that take a considerably less time to resolve[9]. Consequently, The Swedish Courts make a distinction between three types of cases: criminal cases, civil cases and matters (Swedish Courts, 2017). Hence, it seems natural to follow the same distinction with our outputs, in line with what the courts do.

The first category of outputs is resolved criminal cases (BM), which usually takes the longest time to resolve. This measure includes actions punishable by law but also actions concerning property, environmental and planning and building issues.

The second category of outputs is resolved civil cases (TM). Civil cases include actions with a property legal nature such as interpretations of contracts or other financial obligations. Another large group of disputes included in this measure is family law disputes, such as divorces and child custody issues. These types of cases take usually shorter time to resolve, contrary to criminal cases.

The third output measures the number of resolved matters (A). These cases are usually less complicated and can in some instances be decided by staff without legal training.

In summary, the three outputs are resolved criminal cases, resolved civil cases and resolved matters and the four inputs include number of judges, number of law clerks and other personnel and capital measured in office space.

5.3 Descriptive statistics

Table 1 reports all the variables employed in the efficiency analysis and their descriptive statistics for the studied panel, with a five-year interval[10]. The top half of the table shows the main characteristics (mean, minimum, maximum and standard deviation) for the inputs while the bottom half shows the same statistics for the outputs. For the labour inputs, we can observe a rise in the sizes over time. For instance, in 2000, the average number of full-time equivalent judges working in district courts was 6.9 while in 2015 the number increased to 15.8. The same goes for law clerks and the other personnel, the number of employees grew from 5.3 to 15.8 for law clerks and from 13.8 to 28 for other personnel from 2000 to 2015.

One can also observe the large differences in the sizes of the courts. In 2015, the minimum full-time equivalent judges working in district courts were 2, whilst the maximum was 60.7. The variation in district courts sizes are also supported by the large variation in the office space available to the courts. In 2015, for instance, the smallest court measured in office space was only 900 square meters while the largest

[9] Output aggregation could have a substantial effect on parameter estimates and the technical efficiency measures (Kumbhakar, 2015 p. 196).
[10] The five-year interval is used due to lack of space. The mean values of all variables for all years are given in Figure A1 in the Appendix.

was 25 513 square meters. One can similarly observe a gradual increase in office space available to courts over time (Figure A1 in Appendix). This might be due to the merging process, where smaller courts were merged into larger ones with larger office spaces. As mentioned earlier, even though the office spaces grew on average, the total amount of office space used by courts has probably decreased since smaller units have moved to larger ones.

For the outputs, what can be clearly seen in Table 1 is that the number of resolved cases increase with time. Although the caseload is not reported in this table, one can suspect that the increase in resolved cases is due to an increase in the caseload. A sharp increase in resolved criminal cases and in resolved civil cases is visible between 2005 and 2010 (Figure A1 in the Appendix illustrates this). What is also evident is that the number of resolved cases differ substantially for different district courts. For instance, the smallest number of resolved criminal cases for a single court in 2015 is 235.5 while the largest number of resolved criminal cases was 6 970. This is also evident when looking at the standard deviations, which are almost as large as the means.

Table 1. Descriptive statistics for input and output variables for selected years.

	2000			2005			2010			2015		
	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max
Inputs												
Judges	0.9	6.9 (12.5)*	111	0.2	8.51 (12.8)	96	2	12.8 (11.2)	51.8	2.1	15.8 (13.74)	60.7
Law clerks	0.7	5.3 (7.3)	67	0.2	7.35 (8.9)	69	2.7	13.6 (13.1)	67.2	2.74	15.8 (15.3)	77
Other personnel	2.1	13.8 (22.9)	191	0.4	16.93 (27.4)	201	4.9	26 (24.5)	126.2	5.14	28.01 (26.83)	133.95
Office space (square meters)	626	2418.2 (3705.6)	34 448	97	3031 (4572.4)	35 360	949	4267.8 (4176.2)	25 513	900	4511 (4174.8)	25 513
Outputs												
Resolved criminal cases	86	640.93 (925.7)	7949	2	1039 (1396)	1 048 2.5	226.6	1951 (1550.5)	6839.9	235.5	2042.9 (1806.7)	6969.8
Resolved civil cases	50	516.90 (982.1)	8544	2.2	737.6 (1144.7)	8717	175.3	1410.6 (1468.6)	7329.6	183.1	1428 (1513.2)	7147.5
Resolved matters	24	293.93 (544.2)	4930	9.3	443.4 (792.9)	6387.6	100.6	645.9 (611.2)	3737	56	507.9 (513.4)	2955.7
Observations		95			69			48			48	

*Note: Own calculations, data from Swedish National Courts Administration (SNCA). *Standard deviation in parenthesis.*

6 Method

The most common methods used for efficiency estimation are the DEA and the SFA methods (Kruger, 2011). To estimate the efficiency of the judiciary, the literature to this date has tended to solely use the DEA approach, with some exceptions. However, DEA does not allow for hypotheses testing due to its nonparametric-deterministic structure. This issue can be overcome by introducing a two-stage procedure or the bootstrap option. Still, even after performing these techniques the analysis can suffer from severe econometric problems (Simar & Wilson, 2003).

The parametric SFA can handle multiple outputs and multiple inputs models while allowing for hypothesis testing (Coelli et al. 1998). The general form of the stochastic frontier model was first developed by Aigner, Lovell & Schmidt (1977) and by Meeusen and van den Broeck (1977) independently. A panel data version of their model takes the form of

$$\ln y_{it} = X'_{it}\beta + v_{it} - u_{it} \qquad (9)$$

where y_{it} represents the log of the output of court *i* in time *t*, X' is the vector of inputs and β is a vector of parameters to be estimated. The model compromises a two-part error term v_{it} and u_{it}, where the first is a standard noise component and the second represents a non-negative term reflecting technical inefficiency.

In general, technical efficiency panel-data models can be estimated with models that allow technical efficiency to vary over time or be time-fixed[11]. The time-invariant models are those of Schmidt & Sickles (1984) and Pitt & Lee (1981) and are more restrictive in the sense that they assume that courts do not learn over time and never change their efficiency levels. Since a long panel dataset is available for this study (T=17), we want to allow efficiency to vary over time but also to be court-specific. Another reason for letting efficiency to vary over time is that the studied period is characterised by a merging process. Not taking account to the fact that courts may change their efficiency levels due to the merging process will produce inconsistent results.

The standard SFA model as in *(9)* is limited only to one output. This can be solved by using the multi-output distance function, which was previously introduced in the theory section. The technical efficiency estimated will, furthermore, be output oriented: it will analyse by how much a court can increase the level of output while maintaining constant the level of inputs (Kumbhakar & Lovell, 2000). This study will use the stochastic output distance function approach presented in Kumbhakar et

[11] See Cuesta (2000) and Belotti et al. (2013) for a review on time-varying and time-invariant technical efficiency models.

al. (2015). From section 4.2., the general form of the stochastic distance function takes the form of

$$1 = D_0(y_{it}, x_{it}, \beta)h(\varepsilon_{it}); h(\varepsilon_{it}) = \exp(u_{it} + v_{it}) \qquad (10)$$

where v_{it} is the standard error term and u_{it} is the inefficiency term. Imposing linear homogeneity as in *(6)* and using the translog function as in *(7)*, this study will estimate the following model:

$$\begin{aligned}
-\ln y_{N,it} &= \beta_o + \sum_{m=1}^{M} \beta_m \ln x_{m,it} + \sum_{n=1}^{N-1} \gamma_n \ln \tilde{y}_{n,it} + \\
&+ \frac{1}{2}\left[\sum_{m=1}^{M}\sum_{k=1}^{M} \beta_{mk} \ln x_{m,it} \ln x_{k,it} + \sum_{n=1}^{N-1}\sum_{l=1}^{N-1} \gamma_{nl} \ln \tilde{y}_{n,it} \ln \tilde{y}_{l,it}\right] \\
&+ \sum_{m=1}^{M}\sum_{n=1}^{N-1} \rho_{mn} \ln x_{m,it} \ln \tilde{y}_{n,it} + \delta_0 t^2 + \sum_{m=1}^{M} \varphi_m\, t \ln x_{m,it} \\
&+ \sum_{n=1}^{N-1} \varphi_n t \ln \tilde{y}_{n,it} + v_{it} - u_{it},
\end{aligned} \qquad (11)$$

where

$$\tilde{y}_{n,it} = \frac{y_{n,it}}{y_{N,it}}; n = 1, \dots, N-1. \qquad (12)$$

$\tilde{y}_{n,it}$ represents *n*-th output of the *i*-th court at *t*-th year. The corresponding outputs for *n* = 1,2,3 are resolved criminal cases (BM), resolved civil cases (TM) and resolved matters (A), respectively, with resolved matters (A) chosen as numeraire in order to normalise outputs. Likewise, the inputs for *m* = 1,2,3,4 are number of judges (DOM), number of law clerks (NOT), other personnel (AD) and office space (AREA), respectively. As visible in *(11),* except cross-products between the inputs and the outputs, the translog includes products between the time variable (*t*) and the inputs and outputs.

Moreover, the parametric-stochastic models require specification of the distribution of the stochastic parts (Coelli et al. 2005). This analysis will follow a half-normal distribution where

$$\begin{aligned} u_i &\sim i.i.d. N^+(0, \sigma_u^2) \\ v_i &\sim i.i.d. N(0, \sigma_v^2). \end{aligned} \qquad (13)$$

u_i and v_i are distributed independently of each other σ_u^2 and σ_v^2 are the parameters to be estimated together with the parameters in *(11).*

The model will be estimated with a maximum likelihood method. In general, SF analysis is conducted in two steps: first, estimates of the model parameters in *(11)* are obtained by maximising the log-likelihood function. After obtaining the model parameters, point estimates of inefficiency can be obtained through the mean of the conditional distribution $f(u_i|\hat{\varepsilon}_i)$, where $\hat{\varepsilon}_i = y_i - \hat{\alpha} - x'_i\hat{\beta}$. The second step is conducted because the model parameter estimates allow for the estimation of residuals $\hat{\varepsilon}_i$ but not for the inefficiency estimates (Belotti et al. 2013). In order to estimate the efficiency and inefficiency estimates, the well-known solutions by Battese & Coelli (1988) and by Jondrow et al. (1982) will be used. Once the stochastic distance function is estimated, point estimates of the technical efficiencies will be estimated via

$$\text{Technical Efficiency} = \mathbb{E}(\exp(-u_i|\hat{\varepsilon}_i)) \tag{14}$$

as per Battese & Coelli (1988) and technical inefficiencies via

$$\text{Technical Inefficiency} = \mathbb{E}(u_i|\hat{\varepsilon}_i) \tag{15}$$

as per Jondrow et al. (1982). Both estimators of technical (in)efficiency will be presented in the following Results section. This efficiency measure is bounded in the range (0; 1] and represents the relative deviation of the actual output of court *i* from the output it should produce when producing technically efficient on the frontier function. Hence, a measure of 0 means a court is not efficient and a measure of 1 indicates a fully efficient court.

7 Empirical Results

The maximum likelihood estimates of the parameters in the translog stochastic output distance function defined in *(12)* are given in Table 2. The output variables are normalised by resolved matters (A) for imposing the homogeneity restriction. Further, we have interaction terms between the input variables, between the output variables and cross-interactions between the input and the output variables as well between the time variable and the inputs/outputs. One can conclude that most of the parameters are significant at 1%, 5% and 10% level of significance[12]. The loglikelihood is 575 which shows that the model is highly fit. Furthermore, the results given in Table 2 are the first order coefficients and are not easily interpretable as they do not provide much information on the technical efficiency of courts. Based on this, focus will be given to the estimated technical efficiency. The mean efficiency and the mean inefficiency estimates and their standard deviations as per Battese & Coelli (1988) and Jondrow

[12] Also, most of the marginal effect of the parameters at each observation were positive.

Table 2. Estimation of Stochastic Distance Function and Mean Efficiency Scores

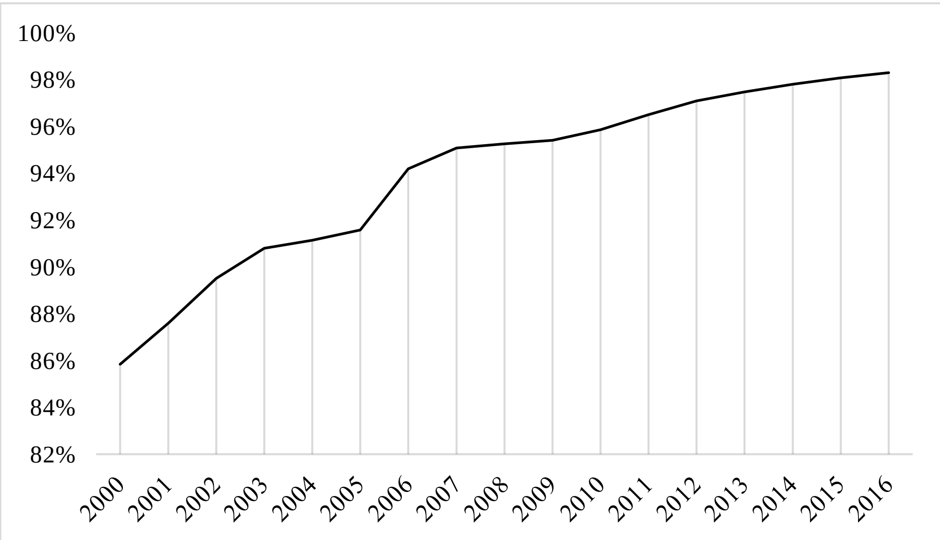
Variable	Estimated Values	Standard Error
Stochastic distance function		
lnDOM	2.353***	0.610
lnNOT	-1.402**	0.667
lnAD	-0.178	0.709
lnAREA	-2.275***	0.378
ln(BM/A)	1.960	2.705
ln(TM/A)	-5.764**	2.925
$(\ln DOM)^2$	0.583***	0.162
(lnDOM)(lnNOT)	-2.265***	0.523
(lnDOM)(lnAD)	1.291**	0.552
(lnDOM)(lnAREA)	-1.447***	0.337
$(\ln NOT)^2$	0.329**	0.167
(lnNOT)(lnAD)	0.304	0.528
(lnNOT)(lnAREA)	0.871**	0.341
$(\ln AD)^2$	-0.376***	0.112
(lnAD)(lnAREA)	-0.247	0.285
$(\ln AREA)^2$	0.278***	0.063
$(\ln(BM/A))^2$	1.931	1.697
(ln(BM/A))(ln(TM/A))	-10.04	6.563
$(\ln(TM/A))^2$	2.945	2.714
(ln(DOM))(ln(BM/A))	0.079	0.418
(ln(DOM))(ln(TM/A))	-0.921**	0.447
(ln(NOT))(ln(BM/A))	0.0766	0.485
(ln(NOT))(ln(TM/A))	-0.268	0.521
(ln(AD))(ln(BM/A))	-0.303	0.499
(ln(AD))(ln(TM/A))	0.735	0.596
(ln(AREA))(ln(BM/A))	-0.464	0.345
(ln(AREA))(ln(TM/A))	1.030***	0.374
t	-0.275***	0.071
tt	0.003***	0.000
t(lnDOM)	0.023***	0.006
t(lnNOT)	-0.017**	0.007
t(lnAD)	0.0005	0.008
t(lnAREA)	-0.009***	0.003
t(ln(BM/A))	-0.021	0.024
t(ln(TM/A))	0.015	0.027
Constant	5.139***	1.673
Efficiency and inefficiency scores		
Mean technical efficiency, via $\mathbb{E}(\exp(-u\|\hat{\varepsilon}_i))$	.93	0.051
Mean technical inefficiency, via $\mathbb{E}(u\|\hat{\varepsilon}_i)$	.075	0.059
Log-likelihood	575.93	
Observations	1,028	

*Source: own calculations, data from SNCA. Note: *** p<0.01, ** p<0.05, * p<0.1* et al. (1982) are presented at bottom of Table 2. The estimated average technical efficiency and inefficiency are 93% and 7.5% respectively for all courts throughout the sample. These estimates demonstrate that district courts are operating on highly efficient scales, but still, with room for improvements.

The model estimated in *(12)* allows furthermore to assess the variations in technical efficiency over time. The estimated mean technical efficiencies from the year 2000 to the year 2016 are plotted in Figure 4. The figure shows that court efficiency has improved steadily over the studied period. For instance, technical efficiency increased from 85.8% in 2000 to 98.3% in 2016. However, we can also observe some periodical variations in the increasing rate of mean technical efficiency scores. Technical efficiency increases substantially between 2005 and 2006, from 91.6% to 94.2%. (Table A1 in the Appendix shows the mean technical efficiency scores for every year in our sample). The explanation for the this might be the large number of court mergers (16 mergers) that happened between the years of 2002 and 2005. These organisational changes might have affected courts performance during 2009, as we observe a level off in the continual growth of technical efficiency [13].

The period between 2010 and 2016 represents the post-reform period as no further merges happened during these years. This is also demonstrated by the steady increase in technical efficiency after 2010, with a peak in 2016 with 98.3%, the last year in our sample. This leads to the conclusion that as courts became larger with time, the efficiency improved. With the given inputs and outputs, technical efficiency is likely to grow until all units reach the frontier i.e. until all courts operate 100% efficiently.

Figure 4. Mean technical efficiency trend for the period 2000-2016



Source: Own calculations, data from SNCA.

[13] One additional plausible explanation for this level off is the economic crisis. The means of the crisis affecting courts efficiency are inconclusive, although one can suspect that the crisis could have produced an increase in the caseload which in turn would have affected courts performance.

Next, since our model allows for estimation of unit-specific efficiencies, the technical efficiency of different courts in the sample will be looked at. Table 3 presents the summary of the results obtained where courts are grouped according to their mean efficiency levels. The second column shows the number of courts in the percentile in absolute terms while the third column as a fraction of total number of courts. Note that the last column indicates the percentage of merged courts in the percentile.

The results in Table 3 shows that 16 district courts, or 15,7% of the total amount of courts had efficiency levels between 100% and 95%. 12.5% of these were district courts that were closed down during the studied period. Moreover, 50% of the courts showed efficiencies between 94% and 90%, from which 35.3% were merged courts. 22.5% and 10.8% of the courts showed technical efficiency between 89 to 85% and between 84% to 80% respectively. All the courts in these percentiles were courts that do not exist as per today. In fact, Table 3 reveals that the majority of the district courts that were shut down are among the courts with the lowest levels of estimated efficiency. This might be the explanation for why we observe lower levels of technical efficiency during the early years of our sample. None of the courts showed efficiency between 79-75% while only one court has estimated efficiency between 74% and 70%.

Table *3. Summary statistics of the Mean technical efficiency scores for the period 2000-2016*

Mean technical efficiency percentiles	Number of District courts		
	In absolute terms	Percentage of total number of courts (%)	Percentage merged courts (%)*
100–95%	16	15.7	12.5
94–90%	51	50	35.3
89–85%	23	22.5	100
84–80%	11	10.8	100
79–75%	0	0	0
74–70%	1	1	100

*Source: own calculations, data from SNCA. Note: *District courts that do not exist as per 2018.*

Taken together, the results show rather large differences in efficiency levels between different courts. The difference between the most efficient and the least efficient court is 23 percentage points. However, these two extremes do not represent courts overall efficiency levels. Figure 5 illustrates that most courts show technical efficiency between 90 and 95%[14]. Hence, the results in this chapter indicate that on average, district courts are largely technically efficient. The next chapter, therefore, moves on to discuss the results implications as well as the main limitations of this study's analysis.

[14]Figure A2 in the Appendix shows the density of the mean efficiency scores.

Figure 5. Mean technical efficiency for period 2000-2016, per court

Source: own calculations, data from SNCA.

8 Discussion

The main findings and their implications as well as limitations of this study will be discussed in the following.

The main finding of this study is that the average technical efficiency of the Swedish district courts is 93% meaning that the estimated average inefficiency lies at 7%. More close inspection of the efficiency levels of each court reveals that none of the courts made a fully efficient use of their resources. Another notable finding is that efficiency in district courts varied significantly across the country. For instance, the difference between the most technically efficient and the least technically efficient court is around 23 percentage points. Additionally, a large difference in effieciency is observed between courts that exist today and merged courts[15]. These results are in agreement with Mattsson's et al. (2018) results which showed that there is a substantial variation in efficiency levels among different courts in Sweden.

Comparison of this study's findings with other studies confirms that Swedish district courts operate on largely efficient scales. Other European countries' first instance courts, for instance, the German, Portuguese, and Italian courts have measured average technical inefficiencies of 11%, 31% and 9% respectively (Schneider, 2005; Santos & Amando, 2014; Favalvigna et al. 2015). This shows that although the efficiency in the Swedish district courts has room for improvements when comparing to other European countries, the district courts in Sweden seem highly technically efficient. However, the aforementioned studies used the DEA method to estimate courts performance and may be therefore inappropriate to compare with this study's

[15] One should bear in mind that the model estimated in (12) does not allow the parameters to differ between merged and non-merged courts. Therefore, the observation made above is strictly based on what Table 7 showed.

results. Besides, cross-country comparisons should be done with caution since different countries face different institutional settings.

Moreover, the first aim of this study was to estimate the technical efficiency of the Swedish district courts using the stochastic distance function approach. In reviewing the literature, no study was found that used this approach when estimating judicial efficiency. The results of the previous chapter demonstrate that the stochastic distance function can very well be used for this purpose. The parametric estimates of the stochastic distance function hold an obvious advantage over DEA estimates since it allows for hypothesis testing. Additionally, it has been shown that when the sample size is relatively large, as in this study's case, the SFA method outperforms DEA (Ruggiero, 1999; van Biesebroek, 2007). Another advantage of the stochastic frontier model is that it has the ability to allow for measurement errors. Measurement errors are specifically important when a long panel data is available. Therefore, it can be concluded that this study's parametric estimates of technical efficiency are more reliable than the non-parametric estimates obtained when using the DEA.

The second aim of this study was to observe technical efficiency over time. Figure 4 showed that the average efficiency levels increased during our studied period. The depicted trend in Figure 4 demonstrates that efficiency levels increased from 85.8% to 98.3%, on average. Moreover, Table 3 reveals that the majority of the district courts that underwent a merging process are among the courts with the lowest levels of estimated efficiency. This might be the explanation for why we observe lower levels of technical efficiencies during the early years of our sample. It may be the case therefore that these variations in efficiency levels across time depend on the merging process.

One source of weakness in this study which could have affected the estimated technical efficiency is that our analysis did not take into account the difference between courts regarding the type of cases and matters they handle. Although a distinction was made between different types of resolved cases, cases can furthermore vary in complexity. For instance, a criminal case concerning theft cannot be compared with a criminal case concerning murder as the two cases require a different amount of resources to resolve. Some courts that showed lower technical efficiency may have encountered more complicated cases, whereas courts that displayed higher efficiency levels may have had less complicated cases. This implication can cause misleading results, why it is important to somehow weight the outputs accordingly. One study that compensates for the complexity of cases is Mattsson et al. (2018) which weights their outputs with hearing time. Unfortunately, this current study could not weight the outputs in the same manner due to lack of data. Hence, further research might explore this possibility.

Another source of weakness in this study is that information on the size of the caseload was not available. The inclusion of the caseload in the analysis would show whether the size of courts' caseloads affects its performance.

Notwithstanding these limitations, our analysis provides important contributions in the research field of courts efficiency. First, this is the first study to use the stochastic distance function approach to estimate the efficiency of courts which helps us to overcome the drawbacks of the much-used DEA method. Second, this study is one of the few studies that observe efficiency levels of courts over such a long period of time. The available detailed data has enabled us to do an in-depth analysis of courts' performance throughout the studied period.

9 Conclusion

The aim of this thesis was to empirically estimate the technical efficiency of the Swedish district courts. This was done with detailed data from the Swedish National Courts Administration using three types of different outputs and four inputs. The efficiency levels were measured using the stochastic distance function approach which has not been used before in this literature. The objective was to observe efficiency levels over time and to observe whether the mergers that happened during the studied period produced increases or decreases in district courts efficiency.

The estimated average technical efficiency for the period between 2000 and 2016 was found to be 93%, while inefficiency was around 7%. A rising trend was observed in the mean technical efficiency and a possible explanation for this might be the merging process that underwent during our studied period. Additionally, the analysis found that there is variation in district courts levels of efficiency. In the future, to develop a full picture of why some courts operate on more efficient scales, additional studies will be needed to study the factors that affect courts performance. Examples of such factors could be the complexity and the size of the caseload, management or judges' education and training.

The policy recommendations here are straightforward. To improve courts performance, it is necessary to look at courts that outperform other courts. Since the analysis concluded that smaller courts displayed lower efficiency levels, one policy conclusion could be that larger units perform better than smaller units. Hence, this study's results provide further support that the decision to merge courts produced positive effects on courts performance.

Taken together, this study offers some valuable insights into the efficiency assessment of courts. However, continued efforts are needed for the expansion of the research concerning the performance of the judicial system.

References

Aigner, D.J., Lovell, C.A.K., & Schmidt, P. (1977), Formulation and Estimation of Stochastic Frontier Production Function Models, *Journal of Econometrics, vol. 6, pp. 21-37*.

Battese, G. E., & T. J. Coelli. (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics 38*: 387–399.

Beenstock, M., & Haitovsky, Y. (2004). Does the appointment of judges increase the output of the judiciary?. International *Review of Law and Economics, 24(3),* 351-369.

Bezat, A. (2011). Estimation of technical efficiency by application of the SFA method for panel data. *Zeszyty Naukowe Szkoły Głównej Gospodarstwa Wiejskiego w Warszawie. Problemy Rolnictwa Światowego, 11(3).*

Castro, M. F., & Guccio, C. (2014). Searching for the source of technical inefficiency in Italian judicial districts: An empirical investigation. *European Journal of Law and Economics, 38(3),* 369-391.

Castro, M. F., & Guccio, C. (2018). Measuring potential efficiency gains from mergers of Italian first instance courts through nonparametric model. *Public Finance Review, 46(1),* 83-116.

Council of Europe European Commission for the efficiency of justice (CEPEJ). (2008). European Judicial System: Efficiency and Quality of Justice. Strasbourg: Technical report.

Coelli, T., & Perelman, S. (2000). Technical efficiency of European railways: a distance function approach. *Applied Economics, 32(15),* 1967-1976.

Cuesta, R. A. (2000). A production model with firm-specific temporal variation in technical inefficiency: with application to Spanish dairy farms. *Journal of Productivity Analysis, 13(2),* 139-158.

Das, A., & Kumbhakar, S. C. (2012). Productivity and efficiency dynamics in Indian banking: An input distance function approach incorporating quality of inputs and outputs. *Journal of Applied Econometrics, 27(2),* 205-234.

Deyneli, F. (2012). Analysis of relationship between efficiency of justice services and salaries of judges with two-stage DEA method. *European Journal of Law and Economics, 34(3),* 477-493.

Falavigna, G., Ippoliti, R., Manello, A., & Ramello, G. B. (2015). Judicial productivity, delay and efficiency: A Directional Distance Function (DDF) approach. *European Journal of Operational Research, 240(2),* 592-601.

Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General), 120*(3), 253-290.

Ferrandino, J. (2014). Testing the Packer Theorem: The Efficiency of Florida's Criminal Circuit Courts. *American Journal of Criminal Justice, 39(2),* 375-393.

Ferrari, A. (2006). The internal market and hospital efficiency: A stochastic distance function approach. *Applied Economics, 38(18),* 2121–2130.

Färe, R., & Primont, D. (1995). *Multi-output production and duality: theory and applications.* Kluwer Academic Publishers.

Hamidi, S. (2016). Measuring efficiency of governmental hospitals in Palestine using stochastic frontier analysis. *Cost Effectiveness and Resource Allocation, 14(1)*, 3.

Irz, X., & Thirtle, C. (2004). Dual technological development in Botswana agriculture: A stochastic input distance function approach. *Journal of Agricultural Economics, 55(3)*, 455–478.

Jappelli, T., Pagano, M., & Bianco, M. (2005). Courts and banks: Effects of judicial enforcement on credit markets. *Journal of Money, Credit and Banking*, 223-244.

Jondrow, J., C. A. K. Lovell, I. S. Materov, & P. Schmidt. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics* 19: 233–238.

Kittelsen, S. A., & Førsund, F. R. (1992). Efficiency analysis of Norwegian district courts. *Journal of Productivity Analysis, 3(3),* 277-306.

Krüger, J. J. (2012). A Monte Carlo study of old and new frontier methods for efficiency measurement. *European Journal of Operational Research*, 222(1), 137-148.

Kumbhakar, S. C., & Lovell, C. K. (2000). *Stochastic frontier analysis.* Cambridge: Cambridge university press.

Lewin, A. Y., Morey, R. C., & Cook, T. J. (1982). Evaluating the administrative efficiency of courts. *Omega, 10(4),* 401-411.

Lovell C.A.K., Travers P., Richardson S., Wood L. (1994). Resources and Functionings: A New View of Inequality in Australia. In: Eichhorn W. (eds) Models and Measurement of Welfare and Inequality. Springer, Berlin, Heidelberg

Löthgren, M., (2000). Specification and estimation of stochastic multiple-output production and technical inefficiency, *Applied Economics, 32:12,* 1533-1540.

Major, W. (2015). Data envelopment analysis as an instrument for measuring the efficiency of courts. *Operations Research and Decisions, 25(4),* 19–34.

Marselli, R., & Vannini, M. (2004). L'Efficienza Tecnica dei Distretti di Corte di Appello Italiani: Aspetti Metodologici, Benchmarking e Arretrato Smaltibile. CRENOS Working Paper.

Mattsson, P., Månsson, J., Andersson, C., & Bonander, F. (2018). A bootstrapped Malmquist index applied to Swedish district courts. *European Journal of Law and Economics*, 1-31.

Meeusen, W., & van Den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International economic review*, 435-444.

Ministry of Justice. (2015), The Swedish judicial system, Government Offices of Sweden, Stockholm.

Nissi, E., & Rapposelli, A. (2010). A data envelopment analysis of Italian courts efficiency. Statistica Applicata – *Italian Journal of Applied Statistics*, 22(2), 199–210.

Santos, S. P., & Amado, C. A. (2014). On the need for reform of the Portuguese judicial system – Does data envelopment analysis assessment support it? . *Omega, 47,* 1-16.

Schmidt, P., & Sickles, R. C. (1984). Production frontiers and panel data. *Journal of Business & Economic Statistics*, 2(4), 367-374.

Schneider, M. R. (2005). Judicial career incentives and court performance: An empirical study of the German labour courts of appeal. *European Journal of Law and Economics*, 20(2), 127–144.

Shepard, R. W. (1953). *Cost and production functions*. Princeton: Princeton University.

Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models ofproduction processes. *Journal of Econometrics*, 136(1), 31–64.

Swedish National Audit Office (SNAO). (2017). *Tingsrätters effektivitet och produktivitet*. Stockholm

Swedish Courts. (2017-03-10). District court. Retrieved from (http://www.domstol.se/Funktioner/English/The-Swedish-courts/District-court/)

Swedish Courts. (2016). *Årsredovisning 2016*. Jönköping

Ruggiero, J. (1999), Efficiency Estimation and Error Decomposition in the Stochastic Frontier Model: A Monte Carlo Analysis, *European Journal of Operational Research, vol. 115, pp*. 555-563.

Tulkens, H. (1993). On FDH efficiency analysis: Some methodological issues and applications to retail banking, courts and urban transit. Journal of Productivity Analysis, 4(1/2), 183-210.

Pedraja-Chaparro, F., & Salinas-Jimenez, J. (1996). An assessment of the efficiency of Spanish Courts using DEA. *Applied Economics, 28(11),* 1391-1403.

Pitt, M. M., & Lee, L. F. (1981). The measurement and sources of technical inefficiency in the Indonesian weaving industry. *Journal of development economics, 9(1)*, 43-64.

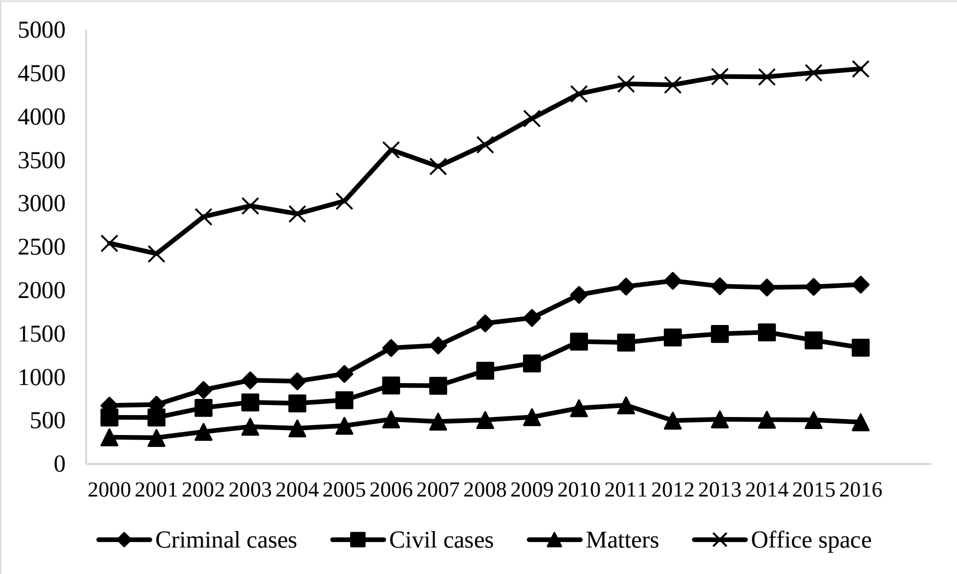
Van Biesebroeck, J. (2007), Robustness of Productivity Estimates, *Journal of Industrial Economics, vol. 55, pp*. 529-569.

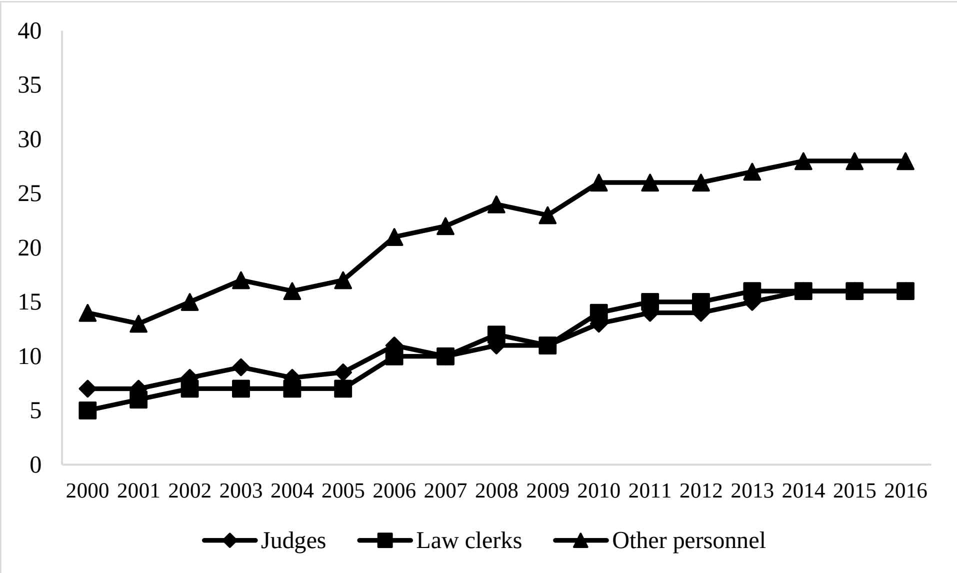
World Bank. (2009). Doing Business in 2009. The World Bank Group: Washington.

Yeung, L. L., & Azevedo, P. F. (2011). Measuring efficiency of Brazilian courts with data envelopment analysis (DEA). *IMA Journal of Management Mathematics*, dpr002.

Appendix

Figure A1. Mean values of inputs and outputs for 2000-2016





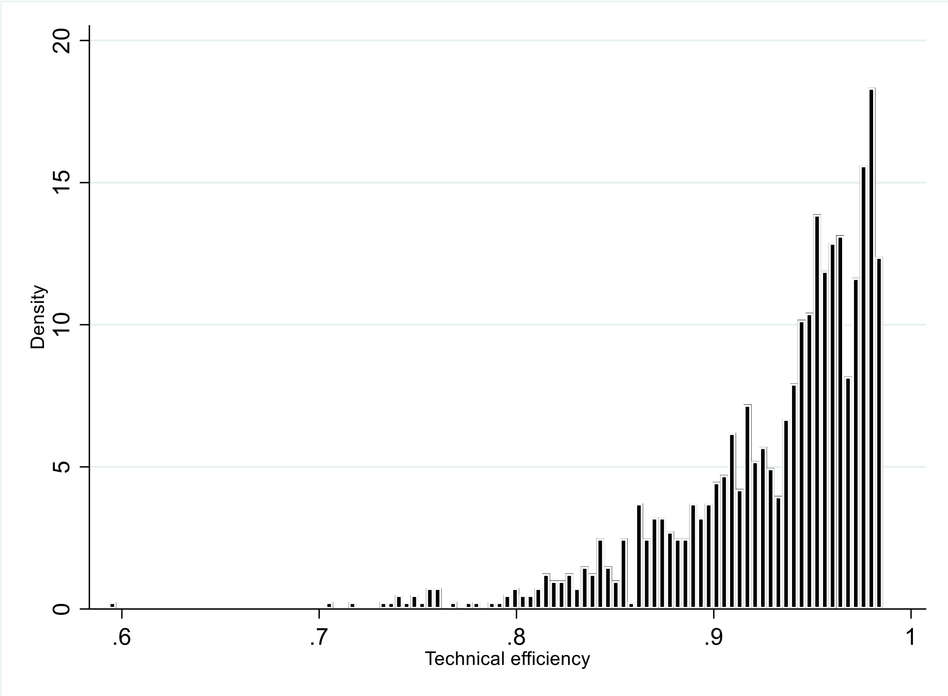
Source: own calculations, data from SNCA.

Table A1. Mean technical efficiency scores per year

Year	Mean	Std. Dev.
2000	.8584	.0614
2001	.8759	.0571
2002	.8951	.0404
2003	.9079	.0311
2004	.9114	.0279
2005	.9157	.0372
2006	.9418	.0151
2007	.9508	.0089
2008	.9525	.0093
2009	.9540	.0080
2010	.9586	.0048
2011	.9650	.0046
2012	.9708	.0038
2013	.9747	.0023
2014	.9779	.0015
2015	.9807	.0014
2016	.9829	.0011
Total	.93	0.051

Source: own calculations, data from SNCA.

Figure A2. Density of the efficiency scores



Source: own calculations, data from SNCA.