Evaluating Canadian Bank Branch Operational Efficiency from Staff Allocation: A DEA Approach

Joseph C. Paradi[1], Elizabeth Min[1] & Xiaopeng Yang[1,*]

[1]Centre for Management of Technology and Entrepreneurship, University of Toronto, Toronto, Canada

*Corresponding author: Centre for Management of Technology and Entrepreneurship, University of Toronto, 200, College Street, Toronto, ON M5S 3E5, Canada. E-mail: xiaopeng.yang@utoronto.ca

Abstract

We examine the operational efficiency of one of the large Canadian banks' branches, which is primarily affected by its strategy of allocating staff and the service quality provided to customers. Two Data Envelopment Analysis (DEA) models are proposed in this research: (1) a staff allocation evaluation model pertinent to employee numbers and transaction volumes, and (2) a customer satisfaction benchmark model to check if the staff allocation scheme meets the expectations of the bank's management. Constant Returns to Scale (CRS) and Variable Returns to Scale (VRS) model results for different branch sizes and geographical regions are presented for analysis. The findings are compared to the bank's current models, validating the use of the proposed DEA models for evaluating operational efficiency from a staff allocation viewpoint in the banking industry. One of the interesting aspects of this work is that the requirement for best practice is not full efficiency but something less. The rationale is that if staff is pushed to the limit, they break and leave – the costs of training and integrating new staff is very high and service levels suffer.

Keywords: *branch; staff allocation; Data Envelopment Analysis (DEA); non-controllable*

1. Introduction

As one of the major institutions close to Canadians' daily lives, banks have a major influence not only on the country's economic development, but also on its entire society. Owing to the increasing number of financial products and customer services aimed at providing convenience and flexibility to their clients, the efficiency of service at the branch is of paramount importance for corporate success. The diverse client base, ranging from individuals, to businesses, large corporations, governments, and non-profit organizations, helps banks grow and deploy their assets. Cumulatively, banks handle approximately 70% of total domestic assets in Canada; the 'Big Five' domestic banks account for over 90% of the assets held by the banking industry and operate through an extensive network that includes over 6,200 branches across Canada (Canadian Bankers Association, 2012). Despite the rapid rise in the use of technology in banking, such as online banking, telephone banking and Automatic Banking Machines (ABM), it was found that, in Canada, 52% of bank customers still conducted transactions with staff at a branch of a financial institution (TNS Canada, 2012). The Canadian Bankers Association found that in 2012, banks employed 275,280 Canadians, and that industry employment, overall, has increased by 14.4% over the past ten years while full-time industry employment has increased by 25.4% over the same period (Canadian Bankers Association, 2012).

To remain competitive customer service is a critical success factor, hence, it is essential to evaluate bank branch operational efficiency; the work here examines several approaches to enlighten management. Optimizing branch staff allocation is still one of the key elements in increasing customer satisfaction, while at the same time reducing costs and increasing productivity. The concept of resource optimization and evaluation is not new; however, the increasingly complex products and services the banks presently offer have made evaluation of the appropriate employee requirements for branches rather difficult. In addition, client-driven activities are usually difficult to predict, adding to the challenge faced by management. Banks traditionally evaluated their performance through different financial measures; however, such conventional methods are often inconclusive and do not reflect the complex banking industry well (Giokas, 2008). Moreover, to evaluate the accuracy of branch staff allocation models, namely how promptly employees respond to customer transaction requirements, most banks are utilizing single ratios

such as quick-response transactions to total transactions. Such a method does not include the impact from staff numbers and fails to capture the multi-dimensionality and complexity of different branch activities (Oral & Yolalan, 1990).

In this study, Data Envelopment Analysis (DEA), a non-parametric multi-dimensional approach, is used as it is one of the widely employed methodologies to evaluate resource optimization for complex business units, such as bank branches. This approach is more effective than traditional methods as they evaluate the branches' relative efficiency against similar units, and identify best performing units to build a frontier for reference instead of just referring to an average value. Among its most significant features is its ability to simultaneously handle multiple indicators of performance as inputs and outputs, hence, it provides an unbiased comparison of similar units without prior specifications of the unknown underlying relationships between these inputs and outputs.

The objective of this study is to develop two DEA models for one of the "Big Five" banks in Canada in order to evaluate its branches' operational efficiency from a staff allocation perspective. The first model evaluates the branches' staff allocation approach to identify best performing branches and the overall effectiveness of the bank's current model for resource allocation. The second evaluates the branches' accuracy of staff allocation in the context of meeting desired customer satisfaction benchmarks set by bank management. Results are compared with the bank's existing models, allowing the models to validate against each other. In this process the analyst discovers improvement opportunities in the models that should lead to better staff resource allocation. The segregated regional and branch size analysis results identify the size characteristics to be employed when adjusting the model to fit different regional and branch size attributes.

The remainder of this paper is organized as follows: Section 2 gives a brief literature review of branch operational efficiency assessment while Section 3 reviews previous studies and presents a detailed introduction to the process of constructing DEA models. Section 4 explores a new approach and discusses its relevance to the application of DEA models in the Canadian banking industry, and various results are compared with traditional bank models. The paper is completed in Section 5, which includes the conclusions and suggests future research directions.

2. Literature Review

Ratio analysis marked the start of performance measurement efforts and is still the most commonly used method across different levels of management and decision-making processes in the banking industry. While ratio analysis is still valid and useful, the rapidly increasing complexity in the banking industry demanded a more sophisticated approach. In response, frontier methodologies have emerged as an important performance measurement approach, allowing more complex use of information to provide insights into performance. Frontier methodologies can generally be categorized into two main approaches, parametric and non-parametric, where the latter is the method we are using in this work.

2.1 Ratio Analysis

Because of its simple form and ease of computation, ratio analysis is one of the most prevalent methods used by the banking sector to measure operational efficiency (Giokas, 2008). The method compares two parameters to capture their relationship, offering insights into different aspects of bank operations, such as profitability, liquidity, asset quality, risk management strategies, etc. Traditional accounting ratios such as return on assets (ROA) and return on equity (ROE) have also been used to measure bank branch efficiency.

Although ratio analysis does provide certain specific insights into performance, it is often incomplete and cannot account for complex operations with its one-dimensional form. Only one aspect at a time can be compared and there is no single aspect of an organization that can fully characterize the operation of a business (Board of Governors of the Federal Reserve System Division of Banking Supervision and Regulation, 2003). Studies have attempted to combine ratios to form a more representative measure of performance in banks and branches (Paradi, Vela, & Yang, 2004). Combining ratios is challenging since it is difficult to determine suitable weights a priori for each efficiency component (ratio), to establish a representative combination (Paradi, Rouatt, & Zhu, 2011). Another problem with ratio analyses is that it is difficult to objectively determine how far above the average is efficient or inefficient (Giokas, 2008).

2.2 Parametric Methods

By virtue of the advantages of allowing for random error and the resulting lower misidentification between inefficiency and noise, parametric methods, such as Stochastic Frontier Analysis (SFA), Distribution Free Approach

(DFA), Thick Frontier Analysis (TFA), etc. have been of interest in efficiency analysis of financial institutions. Nevertheless, such parametric methods are criticized due to the need to impose certain parameters when specifying a functional form (Bauer et al., 1998), which might be inappropriate and thus lead to inaccurate efficiency evaluation results (Greboval, 1999).

SFA has been the most-used parametric method since it was independently introduced in 1977 by Aigner et al., and Meeusen and Van Den Broeck (Reinhard, Lovell, & Thijssenc, 2000; Post, Cherchye, & Kuosmanen, 2002). This method formulates a frontier for single input to multiple outputs or single output to multiple inputs scenarios, assuming a standard normal distribution with a mean of zero and modeling inefficiency using an asymmetric half-normal distribution (Berger, Hunter, & Timme, 1993). However, the half-normal distribution of inefficiency is relatively inflexible and assumes that most units are clustered near full efficiency. Studies including that of Berger and Humphrey (Berger & Humphrey, 1997) have shown that specifying a more general truncated normal distribution for inefficiency yields statistically significant, but different results compared to the half-normal distribution. Despite such increased flexibility, it is still difficult to separate inefficiency from random error, a limitation to this approach.

DFA also specifies a functional form for the frontier, which assumes that random error averages out to zero over time, while efficiency remains stable over time (Bauer et al., 1998). It allows inefficiencies to adopt any distributional shape provided that they remain non-negative. The inefficiency of each unit is calculated as the difference between its average residual and the average residual of a unit on the efficient frontier.

TFA uses the same functional form for the frontier as SFA, but measures the overall efficiency rather than the efficiency of an individual unit, and thus does not assume any distributions for random error or inefficiency. Units in the lowest average-cost quartile are assumed to have above-average efficiency and form a thick frontier, hence the name. Such a property reduces the effect of extreme points in the data; however, it limits the understanding provided of an individual unit's efficiency.

2.3 Non-parametric Methods

As a representative non-parametric method, DEA focuses on comparing the efficiency of a number of Decision Making Units (DMUs), and constructs a rather objective production frontier formed by the dominant DMUs. The most efficient units located on the efficient frontier are those for which no other unit, or linear combination of units, has as much or more of every output (for a set of given inputs) or as little or less of every input (for a set of given outputs) (Charnes, Cooper, & Rhodes, 1978). This condition is also referred to as *Pareto-Koopmans Efficiency* and it defines an efficient DMU where it is not possible to improve any input or output without worsening some of its other inputs or outputs (Cooper, Seiford, & Tone, 2007). DEA differs from its parametric counterparts in that it requires no explicit assumption or knowledge about the relationship between inputs and outputs, and hence DEA does not require any specification of the functional form of the frontier. However, DEA does not account for random error, causing its frontier to be sensitive to the presence of outliers and statistical noise (Bauer, 1990).

There are a few studies that have compared multiple approaches in bank performance evaluation (Knaup & Wagner, 2012; Lee et al., 2011; Wang et al., 2014; Matthews, 2013; Barros, Managi, & Matousek, 2012); however, there is no simple way to determine which of the various methods best evaluate bank performance. The choice of measurement methods appears to strongly affect the calculated efficiency, and results have shown differences in ranking and inefficient unit percentages depending on the method employed. It follows that different methods could offer an advantage in representing the studied relationship depending on the problem at hand.

DEA has shown promising results in bank operational efficiency analysis ever since its introduction and researchers have produced a large and growing body of studies over the last 30 years (Emrouznejad, Parker, & Tavares, 2008; Paradi & Zhu, 2013). DEA determines a comparative ratio of the weighted sum of outputs to the weighted sum of inputs for each unit under evaluation. This relative score expressed as a number between 0 and 1 provides an efficiency measurement comparable to that from parametric methods. Furthermore, DEA's ability to analyze multiple inputs and outputs at the same time, is a strong advantage in evaluating a complex operation such as a bank or a bank branch. These non-parametric properties point to an easier, yet sophisticated, approach to tackle an industry problem, and DEA was judged to be particularly suitable for this study.

2.4 Evaluating Bank Branch Efficiency & Rational Inefficiency

Bank branches have been studied from various aspects including production, profitability, intermediation, cost efficiency, service quality, environment and technology, mergers and acquisition, etc. (Paradi & Zhu, 2013). For all these aspects, the number of employees is generally considered to be an important input affecting the production efficiency of the branch; e.g., the production efficiency of the bank branch is analyzed by Paradi *et al.* (Paradi,

Rouatt, & Zhu, 2011), and the obtained efficiency score is considered to be an input in the second stage model to arrive at the overall operating efficiency of the branch. Instead of using employees, some researchers focused on the service quality of the branch; whereas others have divided staff into more detailed categories such as representatives, tellers, typing, accounting, etc. These categorised staff are then viewed as inputs to produce transactions or financial services (Asmild, Bogetoft, & Hougaard, 2013; Akthera, Fukuyama, & Weber, 2013). But none of the research studied the branch efficiency from both staff allocation and customer satisfaction aspects. This research examines the traditional staff allocation model being used by the bank, and abstracts the pivotal types of staff as inputs to the DEA model to evaluate branch efficiency from the staff allocation aspect.

Bogetoft & Hougaard (2003) discussed *ratioanl inefficiency* in their work as that inefficiency can benefit a firm in the long run. However, they only analyzed inefficency from the viewpoint of capital usage, which emphasized that the firm paying its employees more than their opportunity costs, in order to keep them working efficiently, chould be rational. In fact, rational inefficiency can also be seen from the aspect of relieving employee pressure. The banks measure whether the current staff allocation scheme of bank branches meet their capability requirements, mainly from the view of whether and how customers are satisfied. Specifically, most banks set a target for branches to complete 85% of transactions in a fixed time period, such as 5 or 10 minutes, which ensures that most of the branches are keeping their staff neither excessively idle nor overwhelmingly stressed. The reason why the banks did not set the threshold shorter than 5 minutes or raise their target higher than 85% was precisely because stressed staff would leave and in turn waste more resources. Thus, a relaxed customer satisfaction level can create loyal employees and thereby reduce costly turnover in the labor force (Bogetoft & Hougaard, 2003). On the other hand, rational inefficiency is also reflected in the DEA methodology, as it does not require a branch to be fully optimal in all aspects, but rather one or several merits can make it efficient. This trait of DEA is named sensitivity to extreme values. A more detailed explanation accounting for this will be provided in Section 4.2.

3. Methodology

To introduce a DEA approach for evaluating bank branch staff allocation efficiency, we firstly provide an overview about the models currently used by banks. The bank under study employs a complex staff allocation model based on a queuing algorithm and a single ratio of prompt to total transactions to assess a branch's effectiveness in allocating employees and meeting customer requirements. Based on the understanding of the bank's current models and data characteristics, a further correlation analysis between the results from the bank's method and DEA models are presented to show the weak points in its models, and the validity of DEA in staff allocation analysis.

3.1 Bank's Current Model and Corresponding Benchmark

3.1.1 Bank's Current Model in Staff Allocation

The *Full Time Equivalent* (FTE) measure represents one full time employee's base hours of work in a week, which is 37.5 hours/week. But an FTE can also be made up of two part time staff whose work hours add up to the 37.5 hours per week, so the number of FTEs does not represent the number of different individuals working in the branch. In this study, FTEs are used to measure the number of Customer Service Representative (CSR) units that are responsible for all direct over-the-counter (OTC) transactions in a branch, and are commonly referred to as 'tellers' interacting with customers. The CSR team typically performs three distinct roles: (1) professional services to clients with predominantly transactional banking needs provided by the familiar Common Tellers; (2) personal and business clients' requirements for all their cash handling and transactional banking needs provided by Central Tellers (CT); and (3) complex transactions, such as foreign exchange provided by Client Service Representatives Experts (CSR: Expert).

The CSR resource allocation process involves a complex model as well as management's input to prescribe appropriate staff resource deployment to each branch. Such a process is a cycle where the bank's model uses historical transaction volume data and other model levers to estimate the optimum FTEs per branch by team, which then goes under management's adjustments to add in other factors, such as coverage, to determine the Net FTE by branch. This Net FTE is then used to determine the final maximum number of approved FTEs per branch, and the actual paid FTE data (Paid Data) enters its internal data server with other information including transaction times and the number of transactions. As shown in Figure 1, the cycle is completed as the model uses the updated historical data to re-calculate the optimum resource distribution for the next allocation.
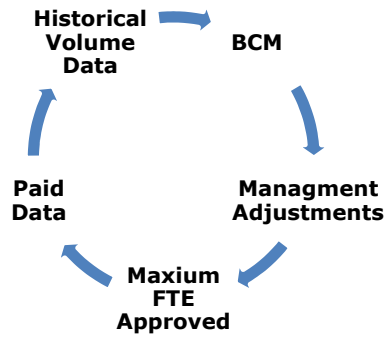
Figure 1. The Bank's CSR Resource Allocation Process

The bank is currently utilizing a commercially available product calibrated to its objectives to determine the optimum FTE counts for CSR teams by branch. Since the bank's model is a commercial software package, it is considered to be a black box model that uses the following model levers and inputs to produce the outputs as shown in Table 1.

Table 1. List of Inputs, Outputs and Model Levers for the Bank's Staff Allocation Model

Model Inputs	Model Levers (Management Inputs)		Model Output
• Historical Volume Data	• Desired Serve Time • Wait Time • Number of Teams • Number of Mapped Assets	⟹	• Required FTEs

The model recommends for each branch the required FTEs by teams according to the previously designated number of teams allocated by corporate management rules. The model uses historical transaction volumes and corporate management's designated inputs, such as desired serve time (minutes), to estimate the transaction volume and recommend the FTE count required to service this volume within a desired transactional time for each branch.

The bank mainly uses Throughput and Client Service ratios to measure branches' CSR allocation efficiencies, which can be broken down into actual paid and recommended ratios, respectively, as shown in Eq. (1-4). Total Paid FTE is the actual paid FTE counts for the branch, and the Net Recommended FTE is the bank recommended FTE counts for the branch. Throughput demonstrates the branch's performance by calculating the average number of transactions completed by one FTE at a branch, and the Client Service ratio demonstrates branch performance by calculating the average percent of time one FTE spends interacting with customers to complete transactions. Total Transaction Time is measured in minutes and it is the total amount of transaction time clocked in a week for the branch.

$$Throughput_{paid} = \frac{Number\ of\ Avg\ Weekly\ Transactions}{Total\ Paid\ FTE} \tag{1}$$

$$Throughput_{rec} = \frac{Number\ of\ Avg\ Weekly\ Transactions}{Net\ Recommended\ FTE} \tag{2}$$

$$Client\ Service\ Ratio_{paid} = \frac{Total\ Transaction\ Time}{Total\ Paid\ FTE \times 37.5 \frac{hours}{FTE} \times \frac{60min}{hour}} \tag{3}$$

$$Client\ Service\ Ratio_{rec} = \frac{Total\ Transaction\ Time}{Net\ Recommended\ FTE \times 37.5 \frac{hours}{FTE} \times \frac{60min}{hour}} \tag{4}$$

The bank uses the above two ratios to compare the actual paid FTE data and the bank recommended FTE data, and

also to indicate the changes over the year. However, such one-dimensional ratio analysis does not clearly demonstrate the staff allocation efficiency of a branch network as aforementioned in the literature review. For instance, all branches require a certain minimum number of FTEs to operate even if all resources are not fully used to meet the transactional volume demand. Such a discrepancy cannot be accounted for in the above ratio analyses and discriminate against smaller branches in comparison to larger branches.

3.1.2 Bank's Benchmark of Staff Allocation

Desired Serve Time (minutes) is the desired average transactional time for the corresponding branch, and its value is decided by corporate management. *Wait time* is a benchmark that the bank aims to meet and is another key model lever used to determine the FTE count. It is based on the percent of transactions that were completed under a certain amount of time. The bank currently aims to complete 85% of transactions under either 5 minutes (85/5) or 10 minutes (85/10) depending on branch characteristics. In fact, 70% of the branches currently operate under the 85/5 benchmark and only smaller branches operate under 85/10 benchmarks. Smaller branches are measured against this more lenient benchmark since all branches require a minimum number of employees to function as a branch and the 85/5 benchmark may be an unreachable standard for smaller branches, which do not require many dedicated representatives as CSRs for their transaction volume.

The bank utilizes a percentage value as the benchmark to measure how well a branch satisfies customers, or in other words, prompt transactions performed by CSRs with respect to the current staff allocation scheme. The branch aims to complete this given percentage of the total transactions under a certain time (5 or 10 minutes) to ensure prompt service and to set a goal for the CSR team to reduce wait time while increasing face time. Such a benchmark can be calculated as follows:

$$Transactions\ Meeting\ Benchmark = \frac{Number\ of\ Transactions\ under\ Benchmark}{Number\ of\ Total\ Transaction} \quad (5)$$

3.2 A DEA Approach

To address the inherent shortcomings mentioned in the preceding sections, we investigate the method of using DEA, a non-parametric tool, to evaluate the complex bank branches' staff allocation efficiency, including the corresponding benchmarks. In this section, the design and definitions of the proposed DEA models, i.e. the inputs, outputs and non-controllable variables are discussed.

Since several input variables encountered in the current study cannot be varied at the discretion of the bank management layer, i.e. branch size, desired serve time, wait time, and number of teams, we employ non-controllable (or non-discretionary) variables, which have been employed in many studies, in input-oriented DEA models (Banker & Morey, 1986; Cooper, Seiford, & Tone, 2007). Assume the data set consists of *n* DMUs, for which **Y** denotes the traditional controllable output variable matrix, and X^C, X^N refer to controllable and non-controllable input variable matrices respectively. The efficiency of DMU$_o$ (the DMU under evaluation) is expressed by the following envelopment form DEA model:

$$\min_{\theta, \lambda} \quad \theta$$
$$s.t. \quad \theta x_o^C - X^C \lambda \geq 0$$
$$x_o^N - X^N \lambda = 0$$
$$y_o - Y\lambda \leq 0 \quad\quad (6)$$
$$\lambda : \begin{cases} \lambda \geq 0\ CRS \\ e\lambda = 1\ VRS \end{cases}$$

in which the input and output vectors of DMU$_o$ are defined by (x_o^C, x_o^N) and y_o. The last constraint is pertinent to the assumed Returns to Scale (RS), where $\lambda \geq 0$ and $e\lambda = 1$ indicate Constant Returns to Scale (CRS) and Variable Returns to Scale (VRS) respectively.

Adding the non-controllable inputs, the production possibility set **P** is defined as follows.

$$\boldsymbol{P} = \{(\boldsymbol{x}^C, \boldsymbol{x}^N, \boldsymbol{y}) \mid \boldsymbol{x}^C \geq \boldsymbol{X}^C \boldsymbol{\lambda}, \boldsymbol{x}^N \geq \boldsymbol{X}^N \boldsymbol{\lambda}, \boldsymbol{y} \leq \boldsymbol{Y}\boldsymbol{\lambda}, \boldsymbol{\lambda} \geq 0\} \tag{7}$$

It can be concluded that the combination ($\boldsymbol{X}^C\boldsymbol{\lambda}$, $\boldsymbol{X}^N\boldsymbol{\lambda}$, $\boldsymbol{Y}\boldsymbol{\lambda}$) formed by vector $\boldsymbol{\lambda}$ outperforms ($\boldsymbol{x}_o{}^C$, $\boldsymbol{x}_o{}^N$, $\boldsymbol{y}_o$) and thus be considered as *efficient*.

The optimal objective function in Eq. (6) is interpreted as the proportionate reduction in all inputs simultaneously required to make DMU$_o$ efficient, and the value of θ is in the range of [0, 1]. $\boldsymbol{s}^{C\text{-}*}$, $\boldsymbol{s}^{N\text{-}*}$ and $\boldsymbol{s}^{+*}$ are the slack vectors in controllable and non-controllable inputs and outputs relative to the efficient frontier. It follows that DMU$_o$ can be improved to be fully efficient by reducing its input excesses and augmenting its output shortfalls using the following scheme:

$$\overset{\wedge}{\boldsymbol{x}}_o{}^C = \theta^* \boldsymbol{x}_o{}^C - \boldsymbol{s}^{C\text{-}*}$$

$$\overset{\wedge}{\boldsymbol{x}}_o{}^N = \theta^* \boldsymbol{x}_o{}^N - \boldsymbol{s}^{N\text{-}*} \tag{8}$$

$$\overset{\wedge}{\boldsymbol{y}}_o = \boldsymbol{y}_o + \boldsymbol{s}^{+*}$$

The new projected DMU defined by Eq. (8) is a point on the efficient frontier and is usually considered to be a target to strive towards in order to achieve desirable and feasible improvements. Furthermore, this target can also be expressed by the positive elements of the optimal solution $\boldsymbol{\lambda}^*$ in Eq. (6) which is conventionally defined as the *reference set* for DMU$_o$.

The input-oriented DEA model incorporating non-controllable input variables, as shown in Eq. (6), is the methodological basis for the current study. Further applications, as well as the results comparison with the bank's traditional models will be introduced in the following section.

4. Application to the Canadian Bank

The collaborating bank studied is one of the "Big Five" Canadian banks, and is currently ranked in the top 100 banks worldwide in terms of asset size (Canadian Bankers Association, 2012). The bank offers an extensive range of financial products and services to customers globally, including personal, commercial and corporate banking, and other financial and investment services. The data used in this study was sourced from the initial pool of over 1,200 branches of the bank. However, this data set was later reduced to 1,166 branches, after eliminating irregularities, missing information, commercial branches and branches without tellers. The remaining dataset includes information on branch characteristics, including market, geographic region, branch size, total number of employees, weekly average transaction and more, as listed in Table 2.

Table 2. List of Data Provided by the Bank on its National Branch Network

Categorical	Numerical	
Branch Background	**Bank's Current Model Data**	Historical Data
• Region	• Number of weekly sessions by branch by team	• Average weekly number of total transactions
• Market	• CSR FTE requirement by team	• Average weekly number of business transactions
• Distribution footprint		
• Branch size	**Model Levers**	• Average transaction time (min)
	• Model serve time	• Paid CSR FTE by team
Hours and Days Availability	• Model wait time	
• Saturday open	• Number of teams	
• Number of days open	• Number of mapped assets	
• Weekday protocol		

4.1 Staff Allocation Analysis

An important step in the efficiency analysis lies in specifying the DEA model, i.e. defining the input and output variables to represent the complex process of allocating staff in actual branch situations. The inputs of the staff allocation DEA model consist of bank FTE counts by team (CSR, CSR: Expert, CT), and the outputs of this model consist of the average weekly number of personal client transactions calculated over the year and the average weekly number of business client transactions calculated over the year (two values that define the customer demand at the branch throughout the year). There are other variables that also affect the use of branch resources, and they are branch size, desired serve time, wait time, and number of teams. These variables were included in the model as non-controllable inputs since they have an effect on the use of resources and producing transactions, but they are not fully controllable by the branch. The following describes the non-controllable inputs used in the model in detail:

- ***Branch Size*** is defined by the total number of employees at the branch.

- ***Desired Serve Time*** ranged from 3 to 5 minutes, and is designated by corporate management.

- ***Wait Time*** is the benchmark number of minutes within which 85% of the transactions were to be completed, and this varied from 5 minutes to 10 minutes.

- ***Number of Teams*** is the number of subdivided CSR teams such as CSR, CSR: Expert and CT present in each branch.

Branch size helps the DEA model group similar branches according to their similar environments. Desired serve time, wait time and number of teams are key model levers for the bank, and thus are designated by corporate management to provide suitable expectations for each branch. Such information is crucial when comparing similar branches with the same expectations. Table 3 summarizes the input and output variables of the DMUs.

With the above setting of 3 normal inputs, 4 non-controllable inputs and 2 outputs, data from 1,166 branches (DMUs) of a Canadian bank are analyzed under CRS and VRS. The calculation results as well as the relevant statistics are shown in Table 4, from which no apparent difference between CRS and VRS results could be found. This is a strong confirmation that the retail bank branches operate on an almost true CRS basis, likely due to the rigidly defined rules and processes under which they must operate.

Table 3. List of Variables for Staff Allocation

Normal Inputs	Non-Controllable Inputs		Outputs
- CSR FTE Counts - CSR: Expert FTE Counts - CT FTE Counts	- Branch Size - Desired Serve Time (min) - Model Wait Time - Number of Teams	⇨	- Average Weekly Personal Transactions - Average Weekly Business Transactions

Table 4. CRS VS. VRS Input-Oriented DEA Model Results for All Branches

Statistics	CRS	VRS
Number of DMUs	1166	1166
Number of Efficient DMUs	154	174
Average Efficiency Score	72.4%	73.4%
Standard Deviation	20.7%	20.9%
Maximum Efficiency Score	100%	100%
Minimum Efficiency Score	18.2%	18.2%

To examine the results found, first we investigate the correlation between Throughput, Client Service ratios and branch size to understand the models used by the bank. Throughput and Client Service ratios do not account for the minimum number of FTEs required to operate a branch and discriminate against smaller branches. This discrimination can be seen from Figures 2 and 3, which plot the Client Service and Throughput ratio values, respectively, against the branch size groups. There is a significant positive correlation between the branch size and the ratios, as the simple regression analysis indicate the medians of the two performance ratios display high correlation coefficients, of 0.92 and 0.77 for the Client Service and Throughput ratios respectively, with branch size. This correlation supports the view that there are weaknesses in using ratio analysis as efficiency measurements for
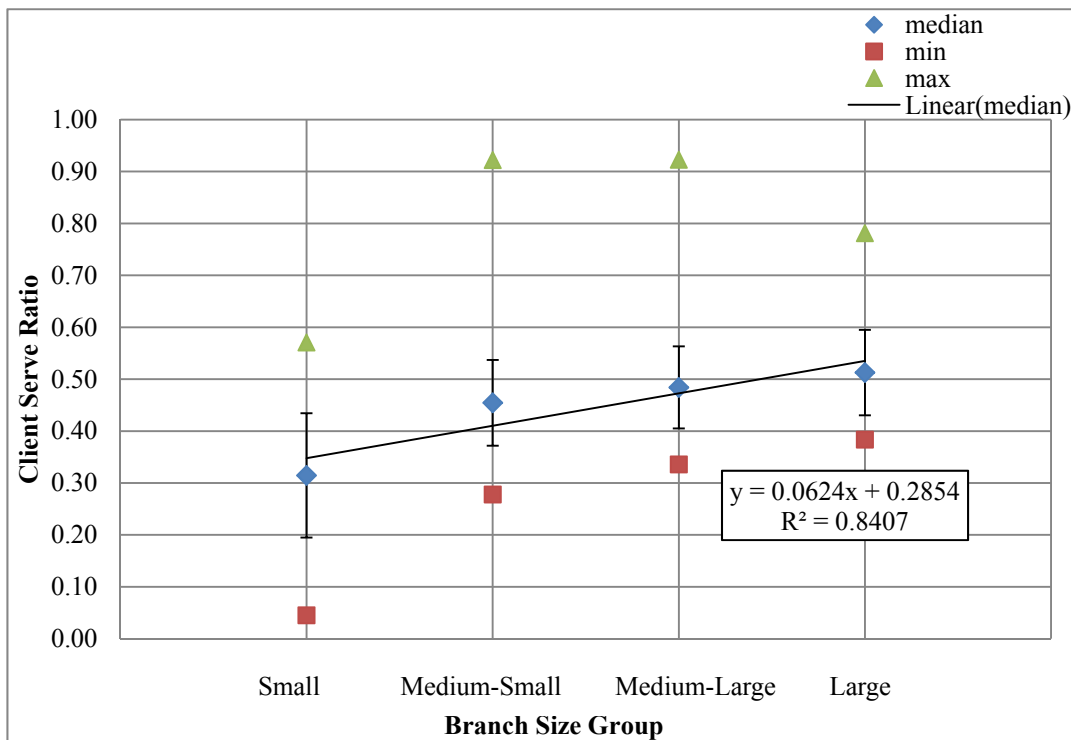
bank branches (Giokas, 2008).

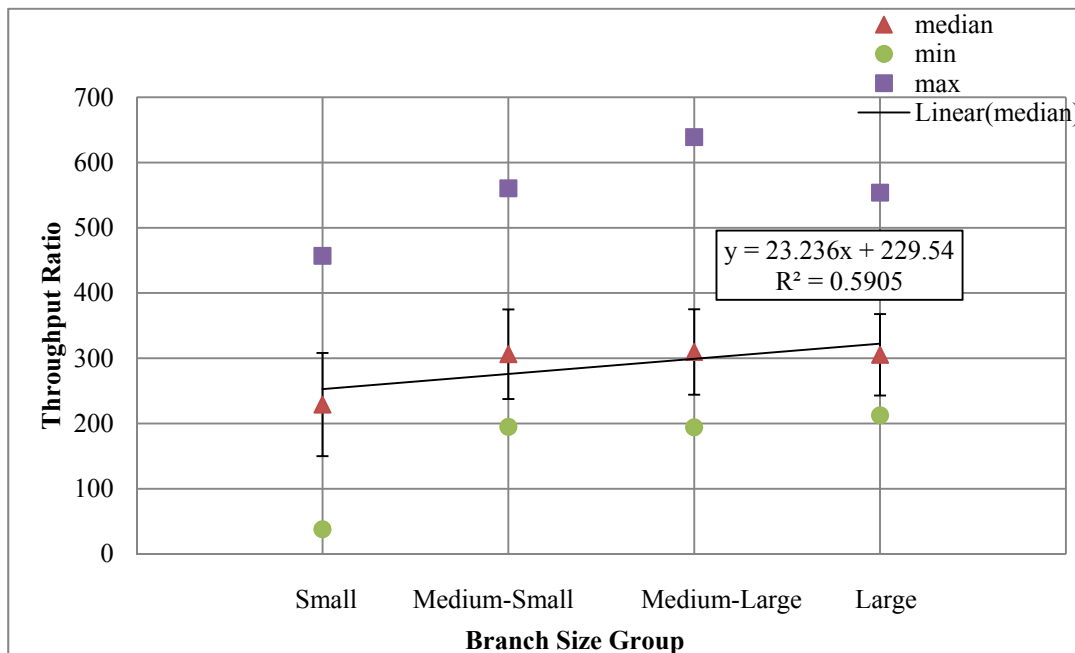**Figure 2.** Distribution of Client Service Ratio by Branch Size Group

Figure 3. Distribution of Throughput Ratio by Branch Size Group

Second, the correlation analysis between CRS DEA efficiency scores (the VRS scores were not examined since there were no apparent difference between CRS and VRS) and the Client Service and Throughput ratios are examined as shown in Figures 4 and 5, respectively. As expected, the CRS DEA efficiency scores did not show significant correlation to the ratios, since no single ratio can appropriately represent the complex resources used in a branch. Ideally, if both results are mapped to the same scale (0, 1), such comparisons should result in a 45 degree line

originating from (0, 0), demonstrating the agreement between the two measurement tools. However, while DEA attempts to incorporate different levels of operations in the model at the same time, the ratios used by the bank are one-dimensional and evaluate only one aspect of the branch operation at a time. Each aspect is an important indicator of a branch's operation; however, it typically does not translate directly into the branch's overall efficiency in staff allocation. The difference in the DEA results and the two performance ratios is a useful indicator for management, and they should consider the use of DEA to identify inefficient and efficient branches to clarify the shortcomings of their current performance measurement ratios.
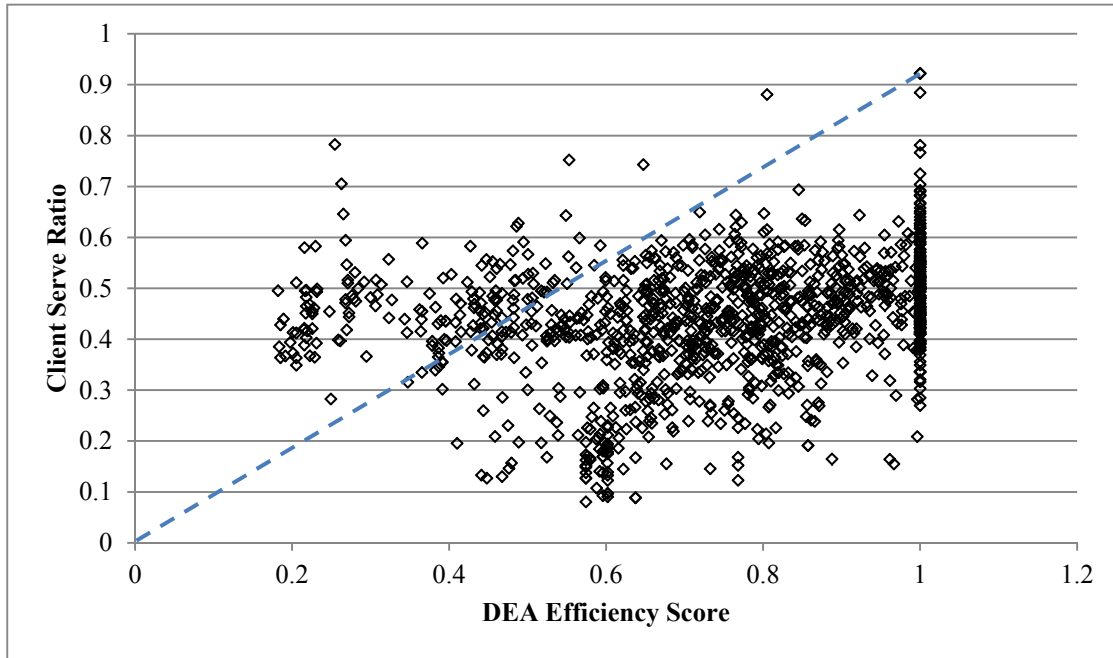

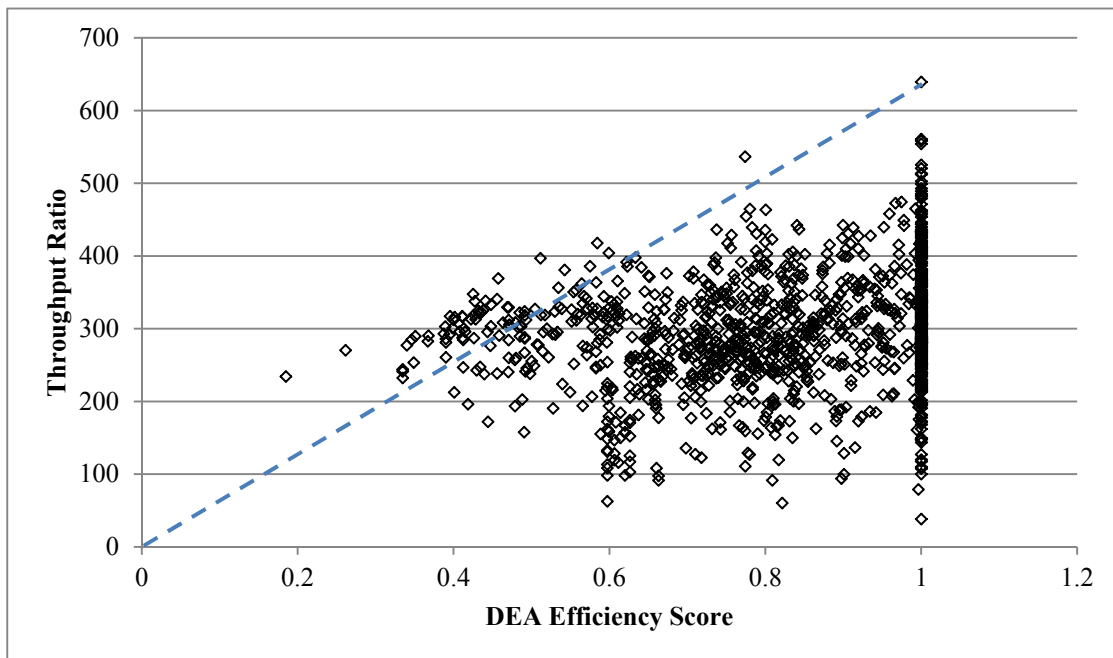**Figure 4.** Comparison of CRS DEA Efficiency Score vs. Client Service Ratio

Figure 5. Comparison of CRS DEA Efficiency score vs. Throughput Ratio

4.2 Customer Satisfaction Benchmark

As explained in Section 3.1.2, the bank has a benchmark measure of performing 85% of all transactions in under either 5 or 10 minutes to reflect Customer Satisfaction by promoting prompt services, and reducing wait times while increasing face time between the customer service representatives and customers. In this section, we provide a DEA approach to evaluate the staff allocation efficiency of the bank under study, in terms of meeting customer satisfaction benchmarks.

We received a month of detailed transactional data for 20 branches, covering either January, 2011 or May, 2011. Since the transaction volume over the year for a bank branch does not change significantly throughout the year, the two different months' data were used together. The branch data was further broken down to an hourly average of transaction data for each branch, increasing the sample size of DMUs to 185.

The normal inputs of the DEA model consisted of management recommended FTE levels and the number of transactions over 5 or 10 minutes (according to the corresponding benchmark) over a one-month period, and the only non-controllable input was desired serve time. The single output variable of the model was defined as the number of transactions under 5 or 10 minutes for the same period. Table 5 displays the input and output variables of the DEA model. At a given total volume of transactions, the main objective of this model is to reduce the number of resources required, as well as the number of transactions exceeding the desired transaction times, while maximizing the number of transactions under the desired transaction time.

Table 5. List of Variables for Customer Satisfaction

Normal Inputs	Non-Controllable Inputs		Output Variables
• Bank recommended FTE • Number of average transactions over 5/10 min for that hour	• Desired serve time	→	• Number of average transactions under 5/10 min for that hour

For the first run, each branch was used as a single DMU. A CRS input oriented DEA model incorporating non-controllable inputs was used, and 7 units were found to be CRS efficient. The average CRS efficiency score was 0.78 with a standard deviation of 0.21, as summarized in Table 6. A high percentage of efficient branches are observed due to the small sample size, exacerbated by the CRS nature of retail banking. This demonstrates one of DEA's major limitations when handling small sample sizes. Therefore the next analysis was conducted with each hour of each branch considered as a DMU, increasing the sample size to 185. Comparison of these two sets of results are given in Table 6.

Table 6. DEA Results of Customer Satisfaction for All Branches

Statistics	Branch as DMU	Hour as DMU
Number of DMUs	20	185
Number of efficient DMUs	7	7
Average Efficiency Score	0.78	0.65
Standard Deviation	0.21	0.20
Maximum Score	1.00	1.00
Minimum Score	0.38	0.31

Figure 6 plots the CRS DEA efficiency score against the bank's Staff Allocation benchmark. When linear regression was performed, the correlation coefficient (r) was 0.70, indicating a significantly high positive correlation between the DEA efficiency score and the bank's internal metric. This indicates that the proposed DEA model and its results are, in fact, a suitable tool to evaluate the bank's staff allocation model with respect to its desired benchmark for customer satisfaction. Often, even when a new technique is demonstrated to be better that an accepted one in an organization, it is rejected because a plausible relationship cannot be demonstrated between them. This is addressed here because the two methods generally agree, even though on an individual basis, better results may be achieved by the DEA model. The correlation indicates a positive relationship between the CRS DEA efficiency score of the proposed model and the bank's benchmark ratio, the portion of bank transactions completed within the benchmark.

Strong correlation can be seen in the branch level data as the correlation coefficient value is close to 1, and thus the proposed DEA model may be preferable for measuring the branches' efficiency according to the bank's desired benchmark.
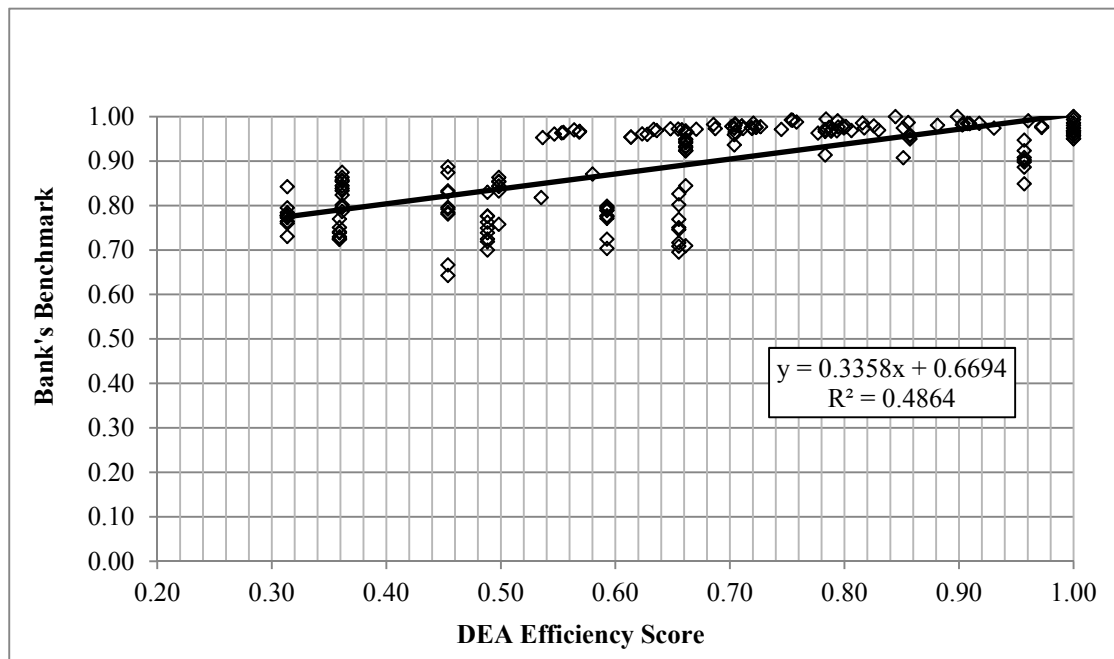

Figure 6. DEA Efficiency Score vs. Bank's Customer Satisfaction Benchmark

From Figure 6, we can also observe that a number of branches deemed efficient by DEA are not efficient by the bank's benchmark. This reflects that DEA is a more tolerant evaluation method for branches which are not leaders in general service quality, but may cost less in manpower usage, may have shorter waiting time, or may have more transactions in a unit of time. DEA perceives such branches as efficient because DEA is sensitive to extreme values, meaning a DMU with either a very low input or a very high output is more likely to be considered efficient. From the management viewpoint, this property of DEA relaxes the requirements for branches to perform perfectly in all ways, which in turn may apply pressure on employees that may result in a decrease in service levels. In this sense, DEA can give a more practical evaluation result that considers the rational inefficiency in separate inputs and outputs (Bogetoft & Hougaard, 2003).

5. Conclusions

Based on the related literature survey in evaluating bank and bank branch efficiency, this research mainly focuses on constructing DEA models for operational efficiency of staff allocation for a major Canadian bank's branch network. A numerical application utilized by the bank for this purpose and the corresponding customer satisfaction benchmarks are also investigated. The comparison between results of the DEA approach and the bank's internal metrics showed that DEA offers a significant potential for evaluating bank branches' staffing efficiency efforts.

The two significant contributions of the current study are, on the one hand, the evaluation of the bank's analysis of Client Service and Throughput Ratios and its appropriateness. By using a DEA based approach, we show that the bank's ratio methods in efficiency evaluation can be improved. We propose a DEA approach in Section 3.2 which incorporates multiple variables, including both controllable and non-controllable measures. Analysis results for various branch sizes and regions from the DEA model were also shown in Table 4. On the other hand, the results for customer satisfaction from the DEA model were highly correlated to the bank's current internal metrics, which lends credibility to the effort of using DEA to evaluate branch efficiency.

Moreover, the factors affecting branch performance may be found by analyzing the differences between efficient and inefficient branches. In-depth observations of the efficient branches that appear frequently and significantly in peer groups for the inefficient branches can provide insights into how to improve inefficient operations. In addition, other environmental factors such as geographic regions, branch size and other model levers used in the bank's models have been shown to indicate characteristic advantages between groups of branches that could be used to calibrate the staff

allocation model to be a better fit. Overall, the DEA results and analysis can be used to estimate the potential savings that could be achieved from the improved performance. This research shows that the bank branches examined have a potential for improvement, which could be achieved by detailed analysis combined with field study.

Further examination and statistical analysis of the DEA results provided in this research could potentially reveal patterns for more effective branch operations, such as the best staff mix between part time and full time staff, best team mix (number of CSR teams) and much more. Also, the limited sample size when evaluating customer satisfaction could be eliminated by obtaining more branch transactional data to produce results reflecting the national branch network, as well as hourly staff allocation efficiency. The proposed models should be extended in the future not only to evaluate the high-level staff allocation models, but also the branch manager level staff scheduling model to closely evaluate the branches' efficiencies and identify more potential savings and improvements.

References

Akthera, S., Fukuyama, H., & Weber, W. L. (2013). Estimating two-stage network Slacks-based inefficiency: An application to Bangladesh banking. *Omega, 41*(1), 88-96. http://dx.doi.org/10.1016/j.omega.2011.02.009

Asmild, M., Bogetoft, P., & Hougaard, J. L. (2013). Rationalising inefficiency: Staff utilisation in branches of a large Canadian bank. *Omega, 41*(1), 80-87. http://dx.doi.org/10.1016/j.omega.2011.01.011

Banker, R. D., & Morey, R. C. (1986). Efficiency analysis for exogenously fixed inputs and outputs. *Operations Research, 34*, 513-521. http://dx.doi.org/10.1287/opre.34.4.513

Barros, C. P., Managi, S., & Matousek, R. (2012). The technical efficiency of the Japanese banks: Non-radial directional performance measurement with undesirable output. *Omega, 40*(1), 1-8. http://dx.doi.org/10.1016/j.omega.2011.02.005

Bauer, P. W. (1990). Recent developments in the econometric estimation of frontiers. *Journal of Econometrics, 46*, 39-56. http://dx.doi.org/10.1016/0304-4076(90)90046-V

Bauer, P. W., Berger, A. N., Ferrier, G. D., & Humphrey, D. B. (1998). Consistency conditions for regulatory analysis of financial institutions: A comparison of frontier efficiency methods. *Journal of Economics and Business, 50*, 85-114. http://dx.doi.org/10.1016/S0148-6195(97)00072-6

Berger, A. N., & Humphrey, D. B. (1997). Efficiency of financial institutions: International survey and directions for future research. *European Journal of Operational Research, 98*, 175-212. http://dx.doi.org/10.1016/S0377-2217(96)00342-6

Berger, A. N., Hunter, W. C., & Timme, S. G. (1993). The efficiency of financial institutions: A review and preview of research past, present, and future. *Journal of Banking and Finance, 17*, 221-249. http://dx.doi.org/10.1016/0378-4266(93)90030-H

Board of Governors of the Federal Reserve System Division of Banking Supervision and Regulation. (2003). *A User's Guide for the Bank Holding Company Performance Report.* Washington D.C.

Bogetoft, P., & Hougaard, J. L. (2003). Rational inefficiencies. *Journal of Productivity Analysis, 20*(3), 243-271. http://dx.doi.org/10.1023/A:1027347616038

Canadian Bankers Association. (2012, October). Retrieved from Canadian Bankers Association: http://www.cba.ca

Canadian Bankers Association. (2012). *Fast Facts about the Canadian Banking System.*

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision- making units. *European Journal of Operational Research, 2*(6), 429-444. http://dx.doi.org/10.1016/0377-2217(78)90138-8

Cooper, W. W., Seiford, L. M., & Tone, K. (2007). *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software.* (2, Ed.) New York: Springer.

Emrouznejad, A., Parker, B. R., & Tavares, G. (2008). Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. *Socio-Economic Planning Sciences, 42*, 151-157. http://dx.doi.org/10.1016/j.seps.2007.07.002

Giokas, I. D. (2008). Assessing the efficiency in operations of a large Greek bank branch network adopting different economic behaviors. *Economic Modelling, 25*(3), 559–574. http://dx.doi.org/10.1016/j.econmod.2007.10.007

Greboval, D. (1999). Managing fishing Capacity: Selected Papers on Underlying concepts and Issues. *FAO Fisheries*

Technical Paper 386, Food and Agriculture Organization of the United Nations, Rome, Appendix XIV.

Knaup, M., & Wagner, W. (2012). A market-based measure of credit portfolio quality and banks' performance during the subprime crisis. *Management Science, 58*(8), 1423-1437. http://dx.doi.org/10.1287/mnsc.1110.1501

Lee, K. P., Cheng, E. T., Yeung, C. A., & Lai, K.-h. (2011). An empirical study of transformational leadership, team performance and service quality in retail banks. *Omega, 39*(6), 690-701. http://dx.doi.org/10.1016/j.omega.2011.02.001

Matthews, K. (2013). Risk management and managerial efficiency in Chinese banks: A network DEA framework. *Omega, 41*(2), 207-215. http://dx.doi.org/10.1016/j.omega.2012.06.003

Oral, M., & Yolalan, R. (1990). An empirical study on measuring operating efficiency and profitability of bank branches. *European Journal of Operational Research, 46*, 282-294. http://dx.doi.org/10.1016/0377-2217(90)90002-S

Paradi, J. C., & Zhu, H. (2013). A survey on bank branch efficiency and performance research with data envelopment analysis. *Omega, 41*(1), 61-79. http://dx.doi.org/10.1016/j.omega.2011.08.010

Paradi, J. C., Rouatt, S., & Zhu, H. (2011). Two-stage evaluation of bank branch efficiency using data envelopment analysis. *Omega, 39*, 99-109. http://dx.doi.org/10.1016/j.omega.2010.04.002

Paradi, J. C., Vela, S., & Yang, Z. (2004). Assessing bank and bank branch performance modeling considerations and approaches. *International Series in Operations Research and Management Science, 71*, 349-394. http://dx.doi.org/10.1007/1-4020-7798-X_13

Post, T., Cherchye, L., & Kuosmanen, T. (2002). Nonparametric efficiency estimation in stochastic environments. *Operations Research, 50*(4), 645-655. http://dx.doi.org/10.1287/opre.50.4.645.2854

Reinhard, S., Lovell, C. A., & Thijssenc, G. J. (2000). Environmental efficiency with multiple environmentally detrimental variables; estimated with SFA and DEA. *European Journal of Operational Research, 121*(2), 287-303. http://dx.doi.org/10.1016/S0377-2217(99)00218-0

TNS Canada. (2012). *report.*

Wang, K., Huang, W., Wu, J., & Liu, Y.-N. (2014). Efficiency measures of the Chinese commercial banking system using an additive two-stage DEA. *Omega, 44*, 5-20. http://dx.doi.org/10.1016/j.omega.2013.09.005