Omega 52 (2015) 33–41

Contents lists available at ScienceDirect

Omega

journal homepage: www.elsevier.com/locate/omega

Evaluation of bank branch growth potential using data envelopment analysis

A.E. LaPlante *, J.C. Paradi

Center for Management of Technology and Entrepreneurship,Faculty of Applied Science and Engineering University of Toronto 200 College Street, Toronto, Ontario, Canada M5S 3E5

ARTICLE INFO

Article history:
Received 7 February 2014
Accepted 24 October 2014
Available online 1 November 2014

Keywords:
Data envelopment analysis (DEA)
Bank branch efficiency
Operations research

ABSTRACT

Banks occasionally employ frontier efficiency analyses to objectively identify best practices within their organizations. Amongst such methods, Data Envelopment Analysis (DEA) was found to be one of the leading approaches. DEA has been successfully applied in many bank branch performance evaluations using traditional intermediation, profitability and production approaches. However, there has been little focus on assessing the growth potential of individual branches.

This research presents five models that examine three perspectives of branch growth. Each model was applied to the branch network of one of Canada's top five banks to gauge the growth potential of individual branches and to provide tailored improvement recommendations. Using various analysis methodologies, the results of each model were examined and their functionality assessed.

1. Introduction

The banking sector is one of the most important threads in the fabric of society in both Canadian and global contexts. As was demonstrated by the 2007-8 economic collapse and ensuing economic difficulties, sound banking practices are essential in maintaining a country's economic health and stability. To ensure more stable economic environments, governments and global committees formulate regulations to govern banking activities. As a result, banks operate under very similar conditions and thus require adept customer service and marketing strategies to maintain or gain market share. In Canada, a near oligopoly exists as there are five very large banks with a total number of over 6,000 branches in Canada and thousands more in other Countries. Their total assets represent 92% of those held by Canadian deposit-taking institutions; and they have combined assets of 3.1 Trillion dollars [29]. In short, they are omnipresent in the life of Canadians.

Even with the rapid advancement of technology, bank branches are still the main conduit through which banks handle more complex transactions and deal with their funds. It is through this extensive network of branches that these banks are able to service their current customers and contact potential clients. It follows that a bank's marketability and growth capabilities are heavily reliant on their branch network as well as the individual growth potential of each branch. To successfully evaluate these criteria, banks must implement performance analysis and target setting at a branch level, rather than at an institutional level. Branch analysis, in many cases, is more desirable and important from a managerial stand point than institutional level analysis. It has the ability to provide information on branch performance that may lead to a better understanding of the variables and relationships that affect a bank's efficiency and profitability [31]. Bank branches are also responsible for a large portion of the value added banking provided to customers and pose the highest operational expenses for a bank. Consequently, cost management can be more effectively performed at the branch level. Continuous improvement of branch performance is crucial in maintaining a competitive standing in the financial industry.

Despite the many attempts to accurately measure bank branch efficiency, the multi-faceted nature of bank branches and the complexity of the services they provide have made this a difficult task [26]. Although banks employ many performance analysis techniques of their own, many of them produce conflicting results or lack the robustness required to fully appreciate the intricacy of the relationships that exist [31]. Financial ratios are unable to simultaneously consider all variables and are, therefore, unable to comprehensively describe branch performance. On the other hand, traditional profitability measures do offer some desirable characteristics, but comparative analysis of branches can be misleading. Their use of averages can make the identification of and

* Corresponding author. Tel.: + 416 978 6924x1; fax: 416 978 3877.
E-mail address: alex.laplante@mail.utoronto.ca (A.E. LaPlante).

http://dx.doi.org/10.1016/j.omega.2014.10.009
0305-0483/

comparison to top performers very difficult. Aside from the numerous profitability measures and financial ratios, banks use frontier efficiency analysis to objectively identify best practices within their organizations [38]. Amongst the frontier efficiency analyses acknowledged in literature, Data Envelopment Analysis (DEA) was found to be one of the most versatile approaches used in the banking industry [32].

This study focuses on one of Canada's major national banks, henceforth referred to as the “Bank”, which employs various performance measures and statistical analyses to evaluate its branch network performance and to assess the growth potential of each branch. Through its analyses, the Bank attempts to determine the best available methods to gain market share and increase its Share of Wallet (SOW). The Bank uses the results from these analyses to determine resource allocation in branches and to help focus and implement marketing strategies. These results are also used in determining the location of new branches, as well as branch relocations and closures.

In planning to be able to better evaluate the growth potential of each branch in the Bank's branch network, a branch network Data Envelopment Analysis (DEA) was proposed. This analysis involved several DEA models, whose individual foci include relative market efficiency, branch churn efficiency, and year to year branch growth. Hence, the aim of this research is to provide the Bank with a more comprehensive approach to measure bank branch efficiency that identifies the growth potential of their branches, an aspect of banking not covered by traditional production, profitability and intermediation models. In doing so, this article will attempt to bridge the gap that exists in current DEA literature with respect to measuring growth directly from the results of a DEA model without the use of Malmquist indices.

The remainder of this article is organized as follows: Section 2 briefly reviews the literature on DEA used in bank branch efficiency analysis; Section 3 discusses the DEA methodologies and data used in the study; Section 4 focuses on the DEA models and their objectives; Section 5 reports on the main results of the empirical tests; and the main conclusions are presented in Section 6.

2. DEA in bank branch efficiency analysis

Since its conception, DEA has become one of the most widely used approaches to measure the efficiency of financial institutions. However, the majority of DEA banking studies have focused on banks at an institutional level, rather than at the branch level. This can be partially attributed to the difference in data availability. The majority of banks are publicly traded firms that are listed on major stock exchanges and thus, must provide their investors with quarterly and annual financial reports. This makes the collection of data for institutional level analyses rather easy. On the contrary, branch level data is proprietary information and is not generally disclosed to the public. Instead, it is either aggregated into bank financial reports or not reported at all. Nonetheless, surveys have shown that there has been a steady increase in DEA branch studies, nearly doubling in the last five years alone [32].

To date, there are four survey papers that review DEA applications in the banking industry, of which three focus on firm level applications and only one on branch level applications. The first to review the major efficiency techniques used in the evaluation of bank performance were Berger and Humphrey [10]. This survey reviewed a total of 103 papers including 57 DEA based papers. Of these, 42 focused on bank level analysis while the remaining 15 focused on branch level analysis. Berger [9] provided a review and critique of over 100 studies that compared cross-national bank efficiencies obtained using various frontier techniques. Fethi and Pasiouras [21] presented a review of 196 studies which employed operational research and artificial intelligence techniques to assess bank performance. Of these studies, 151 used DEA or similar techniques and 30 focused on evaluating efficiency at the branch level. Most recently, Paradi and Zhu [32] published a survey that focused heavily on the use of DEA in branch analysis. Among the 285 bank related DEA publications identified, 90 focused on branch analysis and were discussed in greater detail.

Branch level DEA applications can have a diverse set of business objectives; however, the majority of these reports focus on evaluating branch specific operations. These studies allow for the exploration of efficiency determinants and provide the capability of identifying deficiencies in areas that are controllable by branch managers. That being said, branch performance measurement is not a simple task. Branches come in an assortment of sizes, operate in different economic regions and offer a variety of services to a diverse range of customers. In order for a branch performance analysis to be significant and reliable, it should capture the critical aspects of the bank's internal operating processes, leading to a more adept understanding of these processes. Moreover, the analysis should provide target setting through the identification of best- and worst-practices and offer the capability of investigating the sources of the inefficiencies.

Depending on the objective of the analysis, different DEA model frameworks exist. Of these models, there are three that are very commonly used in branch analysis; intermediation ([15,18,5,3,17]), production ([36,33,23,35,6,12,34,24,17,37,31]), and profitability ([27,1,28,30,31,37]). Additionally, the market model ([4,27,1,28,30,31,37]), employed in this study, is also occasionally used.

Measuring bank branch growth directly from a DEA model has yet to be achieved in literature. There are, however, DEA papers which employ Malmquist indices to measure the Total Factor Productivity (TFP), technical change and efficiency catch up of each branch. Camanho and Dyson [11] used Malmquist indices to evaluate group performance reflecting the impact of environmental factors and regional managerial policies. Gaganis et al. [22] employed Malmquist indices to analyze the efficiency and productivity of branches, while Asmild and Tam [2] estimated global frontier shifts of branch networks using a global Malmquist measure.

3. DEA methodologies and data source

3.1. Data envelopment analysis

Data Envelopment Analysis (DEA), first introduced by Charnes, Cooper and Rhodes in 1978 [14] extended Farrell's concept [20] of estimating technical efficiency through the comparison of each organizational unit with the efficient Production Frontier. DEA is a non-parametric linear programming technique that defines the set of best-practice or frontier observations as those for which no other Decision Making Units (DMU) or linear combination of DMUs has as much or more of every output for as much or less of every input. DEA produces a convex production possibilities set by connecting the best-practice observations with a piecewise linear frontier. All best-practice DMUs sitting on the frontier are considered efficient and receive an efficiency score of 1. The units not on the frontier are considered inefficient and receive an efficiency score of less than 1. These scores are calculated by projecting the inefficient unit onto the efficient frontier.

The DEA methodology offers a number of advantages over traditional parametric techniques. It makes each DMU look as favorable as possible to its peers by allowing each DMU to choose its own variable weights/multipliers. This characteristic makes DEA an ideal choice when assigning numerical values to variables proves to be difficult or when variables are qualitative in nature. DEA also has the ability to identify reference units for each DMU, which proves to be a very useful managerial tool as it aids in determining the potential causes and remedies for the identified inefficiencies [19]. Additionally, DEA does not require prior assumptions of the observation's distributions.

Since the conception of the Constant Returns to Scale (CRS) model [14], DEA has been subject to numerous theoretical advances and methodological extensions. Perhaps most notably were the development of the BCC model by Banker et al. [7], which allowed for variable returns to scale (VRS) and the Slack-Based Model (SBM), which is unit invariant and has an efficiency measure that is monotone decreasing in the slacks of the input and output variables [16].

In this study, both input and output oriented VRS models are used to evaluate the relative performance and growth potential of bank branches in order to take advantage of their translation and unit invariant properties. The general formulation of the Input Oriented VRS model is provided below. Further detail concerning model choice will be provided in Section 4.

Primal VRS Input-Oriented Model (1)	Dual VRS Input-Oriented Model (2)
Maximize:	Minimize:
$w_0 = \sum_{r=1}^{s} u_r y_{r0} - \tilde{u}_o$	$z_0 = \theta - \varepsilon\left[\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right]$
Subject to	Subject to
$\sum_{i=1}^{m} v_i x_{i0} = 1$	$0 = \theta x_{i0} - \sum_{j=1}^{n} x_{ij}\lambda_j - s_i^-$
$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} - \tilde{u}_o \le 0$	$y_{r0} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+$
$-u_r \le -\varepsilon$	$1 = \sum \lambda_j$
$-v_i \le -\varepsilon$	$0 \le \lambda_j$, for j=1,…,n
$\tilde{u}_0$: *free in sign*	$0 \le s_i^-$, for i=1,…,m
	$0 \le s_r^+$, for r=1,…,s

The θ variable is the input reduction applied to the DMU under consideration. This decrease is applied to all of the DMU's inputs in order to bring it closer to the frontier. The θ variable is constrained between zero and one, with one representing full efficiency or radial efficiency. This variable also denotes the required percent of inputs inefficient DMUs would be required to reduce in order to theoretically produce the same amount of outputs. The value of the i^{th} input to unit j is represented by x_{ij}, while the value of the r^{th} output from unit j is represented by y_{rj}. The non-negative intensity variables, λ, represent the weight of each of the n DMUs.

Once the Dual model has calculated the radial efficiency of the DMU it will then calculate the input and output slack variables, also referred to as input excesses, s_i^-, and output shortfalls, s_r^+. If slacks are present, the DMU is considered to have mix inefficiencies. Thus, it will require further reductions beyond the optimal θ^* reductions found in stage one and the input proportions will need to be adjusted. If both slacks are zero and θ^* is equal to one, then the DMU is considered fully efficient. Consequently, if η^* is equal to one with non-zero slacks, then the DMU is considered radially efficient with mix inefficiencies, or weakly efficient.

In order to take into consideration exogenously fixed variables which are pertinent to the model but beyond the discretionary control of the DMU's manager, Banker and Morey extended the mathematical formulations introduced above [8]. These models took the following form where 'D' is the subset of discretionary variables while 'ND' is the subset of Non-Discretionary variables.

Primal VRS Input-Oriented Model (3)	Dual VRS Input-Oriented Model (4)
Maximize: $w_0 = \frac{\sum_{r=1}^{s} u_r y_{r0} - \tilde{u}_o}{\sum_{i=1}^{m} v_i x_{i0}}$	Minimize:
	$z_0 = \theta - \varepsilon\left[\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right]$
Subject to	Subject to
$\frac{\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} - \tilde{u}_o}{\sum_{i=1}^{m} v_i x_{ij}} \le 1, \quad j \in \{1,\ldots,N\}$	$0 = \theta x_{i0} - \sum_{j=1}^{n} x_{ij}\lambda_j - s_i^-, \quad i \in D$
$-v_i, -u_r \le -\varepsilon$ for $r \in \{1,\ldots s\}$ and $i \in D$	$0 = \theta x_{i0} - \sum_{j=1}^{n} x_{ij}\lambda_j - s_i^-, \quad i \in ND$
$-v_i \le -\varepsilon$ for $i \in ND$	$y_{r0} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+$
$\tilde{u}_0$: *free in sign*	$1 = \sum \lambda_j$
	$0 \le \lambda_j$, for j=1,…,n
	$0 \le s_i^-$, for i=1,…,m
	$0 \le s_r^+$, for r=1,…,s

3.2. Data overview and treatment

The collaborating bank in this study is one of the "Big Five" Canadian banks and within the top 75 banks worldwide in terms of asset size [13]. The Bank includes a branch network of more than 1, 000 branches and employs over 40,000 branch and corporate personnel. The Bank offers retail, commercial and corporate banking and an extensive range of financial products and services. The following table (Table 3.1) provides a partial list of products and services offered by the Bank. These services are administered through several channels including in-branch, debit cards, Automated Banking Machines (ABMs), and internet and telephone banking.

The Bank provided a substantial data set, with upwards of 70 variables, including categorical (region, market etc.) and non-categorical (products, assets etc.) variables for over 1000 of its branches; offering the opportunity to evaluate the problem at hand from several vantage points through the construction of multiple unique models. Although issues with dimensionality were unlikely with such a large number of DMUs, DEA loses discriminatory power as the dimensionality of the production space increases. Therefore, to obtain more informative and distinct results, careful variable selection was performed by means of correlation analysis and principal component analysis (PCA). It should be noted, that the results of these tests were carefully considered along with the overall goal of the model and the desires of bank management in mind. There are some cases where the removal of variables does not make sense from a managerial stand point although it does from a mathematical point of view [25].

In addition to the Bank provided variables, several indices were created to account for local demographics and surrounding competitor density. GPS coordinates for each of the Bank's branches along with those of all competing bank branches were collected from the Pitney Bowes database available through the University of Toronto Library system. Using these GPS coordinates, each branch was matched to its closest dissemination areas (DA)[1]. Average household income and population data was collected for each DA in order to create real income and population indices measured in volumes. A competitive index was also created by determining the number of competitor branches within 25, 10, 5 and 1km of each of the Bank's branches. These three indices were incorporated into each model to account for the effect of local environment and market conditions.

[1] A dissemination area (DA) is a small, relatively stable geographic unit composed of one or more adjacent dissemination blocks. It is the smallest standard geographic area for which all census data are disseminated. DAs cover all the territory of Canada.

Table 3.1
Bank's retail and commercial products and services.

Retail and commercial products and services	
• Bank accounts (chequing, savings) • Lines of credit • Mortgages and other loans • Mutual funds and Investment services	• Investment banking and brokerage • Credit card • Foreign exchange, wire transfers and bank drafts • Insurance brokerage

Summary statistics for the variables used in each model are provided in Appendix A.

For comparison purposes, the data was also segmented by select variables. Categorical variables were segmented by the Bank's predefined categories (Central/Western/Eastern Canada; Urban/Rural) while non-categorical local demographic indices (population and income) were segmented using k-means segmentation.

4. Study focus and DEA modeling

4.1. Study focus

Like any business or corporation, the Bank continually strives to improve their growth, profit margins and increase their market share in order to remain competitive. In the current economic times, the retention and growth of customers and funds has become increasingly difficult and is a major focus of the Bank. It was therefore decided that the focus of this study would be to examine the growth potential of the Bank's branches and identify the best performers using DEA. The model presented in this study each offer a unique perspective and direction for attaining their shared goal. In total, three model perspectives and five models, each using a unique combination of inputs and outputs, were defined.

4.2. Lost accounts/gained accounts model

Banks, along with many other service industries, take great interest in understanding the flux of clients, accounts and services. The retention and development of existing clients and the acquisition of new clients each play a vital role in the market growth and overall health of a bank. Using variables that defined attrition, existing client growth and new client gains over the 2010–2011 period, the “Lost Accounts/Gained Accounts” model looks to assess which branches were best at retaining and growing their customer base, funds and products. A branch that retains customers, attracts new customers and grows the product use of existing customers will have a higher efficiency.

The model was constructed with the reasoning that a best performer should minimize attrition (inputs) and maximize growth and client gain (outputs). With this approach in mind, the following model was developed: Table 4.2.

Attracting new clients is the usual goal of corporate marketing campaigns and location, both of which are out of the hands of the branches themselves. Minimizing the attrition of current clients, however, can be more easily managed at the branch level. It follows that an input oriented model was seen to be most suitable for this model. The input oriented VRS model is also translation invariant with respect to the output variables which proves to be a very useful attribute as the number of retained customers and their growth in the products could have negative values. The income, population and competitive indices were treated as environmental, hence non-controllable variables.

4.3. Market models

Although the market model is relatively uncommon among bank branch efficiency analyses, it provides the opportunity to examine how branches are performing relative to their local market. The market model measures the extent to which a bank branch, given its capacity and available resources, realizes its potential to sell products and provide services in a given market. To achieve market efficiency, a bank branch must expand its outputs and optimize how cohesively the branch (size, number of employees, etc.) fits into its market conditions. Unlike the other models contained in this study, market models have been occasionally used in literature. Athanassoupoulos' [6] study is, however, the only one with published results.

The Market models developed in this study examine a branch's ability to use its resources and local market conditions to grow its market share and produce new accounts. Input variables for this model were defined as resources and uncontrollable environmental variables which included branch characteristics and local market conditions, while outputs included all branch products and funds. For the market models, an output orientation was chosen as the input variables are largely uncontrollable by management and the ultimate goal of the model is to maximize outputs given a branch's specific market condition and resources. It follows that if a branch was operating with fewer resources or in worse market conditions (higher levels of competition) and was able to produce the same amount of funds as a branch with more resources or better market conditions, then the first branch would be considered more efficient. The finalized Market model can be viewed in Table 4.3.

Combining variables to improve the discriminatory power and reduce the dimensionality of a DEA model is a common practice used in the literature. To investigate the impact of reducing the number of variables, the individual products (i.e. Total deposits, investments, RRIFs, RRSPs and lending) were replaced with an aggregated ‘funds managed’ metric (equal to the sum of all products). The reduced variable model is shown in Table 4.4. To maintain clarity, the original model is referred to as the “Component” Market model, while the reduced variable model is referred to as the “Aggregate” Market model.

Table 4.2
Lost accounts/Gained account model.

Input	Output
Attrited product count	Retained customers
Attrited customers	Retained customer growth in product count
Attrited funds managed	New customers
Local household income[a]	New funds managed
Local population index[a]	#Competing branches in a 10 km[a] radius

[a] non-discretionary environmental variables

Table 4.3
Market model- components.

Input	Output
# of customers	Total lending balance
Total # of most valuable customers	Total investment balance
Number of employees	Total RRSPs
Local household income index[a]	Total RRIFs
Local population index[a]	Total deposits
	#Competing branches in a 5 km radius [a]

[a] non-discretionary environmental variables

Table 4.4
Market model- aggregate.

Input	Output
# of customers	Product count
Total # of most valuable customers	Total funds managed
Number of employees	Branch's share of local market
Local household income index[a]	Growth in the branch's funds managed
Local population index[a]	Funds managed: share of local market
	#Competing branches in 5 KM[a]

[a] non-discretionary environmental variables

4.4. Delta models

These models employ delta values, calculated using several of the previously introduced variables, to assess a branch's ability to grow customers, funds and products. The Bank provided variables that measured the change in the existing product count and funds between 2010 and 2011. This enabled the construction of a model that considers overall growth as well as existing client growth.

Using the 2010 and 2011 data, the delta values (Δ value = 2011 value - 2010 value) were calculated for the number of customers, product count and funds managed. Delta values for the growth in product count and funds managed of existing customers were provided by the Bank. The final delta model is presented in Table 4.5.

A second Delta model that used the percent change of each variable was also constructed (Percent Change = 2011value/2010value) and the results of the two models were compared. For clarity's sake, the model using delta values is referred to as the “Difference” model, while the model using percentage change is referred to as the “Percent” model.

For similar reasoning to that provided for the Lost Accounts/Gained Accounts model, an input oriented VRS model was implemented. Again, this model choice also provides the benefit of translation invariance in its outputs.

4.5. Local analysis

Along with global analyses, a local analysis was also performed for each model. Using k-means segmentation, the DMUs were segmented based on their local average household income and population indices. To determine the number of clusters, the within cluster sum of squares was calculated for a variable number of clusters and plotted against the number of clusters. Fig. 4.1.

From this plot, it was estimated that five or six clusters would be ideal. Ultimately, five clusters were chosen. These are summarized in Table 4.6.

Table 4.5
Delta model.

Input	Output
Number of employees	ΔCustomers
Local household income index[a]	ΔProduct count
Local population index[a]	ΔFunds managed
	#Competing branches in 10 KM[a]

[a] non-discretionary environmental variables

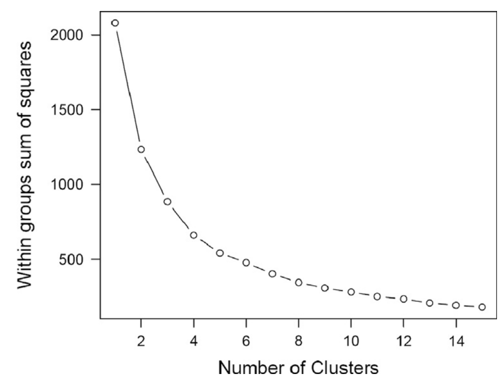

Fig. 4.1. Within group sum-of-squares vs. Number of clusters.

Table 4.6
Cluster statistics for principal component analysis.

Cluster	Number of DMUs	Mean population index	Mean income index
1	45	710.02	168,288.29
2	542	641.09	89,695.87
3	43	548.00	55,541.57
4	217	700.13	116,119.83
5	193	577.27	70.910

5. Results and discussion

5.1. Principal component analysis

Principal component analysis (PCA) is used to evaluate the significance of a variable in a model by comparing the efficiency distributions obtained by running the model with and without the variable by means of a Wilcoxon rank sum test [16]. The principal component analyses performed in this study provided insights into the relationships that exist between certain variables and indices and each type of branch efficiency. Table 5.1.

Some significant findings can be noted from the summarized results:

- Contrary to initial hypotheses, the total number of competitors found in a 1 KM radius seem to have an insignificant impact on branch efficiency while the number of competitors in a 5 KM and 10 KM radius have consistent statistically significant impacts on most of the calculated efficiencies. This may be attributed to the fact that Canada is a vast country and aside from the large cities, citizens generally live in suburban areas where businesses are more spread out. It follows that the number of competitors in a broader area has more impact on a branch's efficiency than the number of competitors immediately surrounding the branch. It would be interesting to investigate if this holds true for branches exclusively located in metropolitan regions.
- The average household income of the branch's surrounding area was found to be statistically significant in all models. Conversely, the local population index was found to be insignificant when used in the Delta model. This may indicate that although population affects the number of clients, products and funds a branch can initially attract, it does not necessarily significantly contribute to the growth in these variables from one year to the next.
- The market model initially included several additional environmental variables not included in the other models. The number of ABMs located within a branch was found to have an insignificant impact on the market efficiency. Similarly, the number of growable clients (a Bank metric denoting clients who hold a large amount of funds external to the Bank) had no impact. It was found that the number of clients who held a large number of funds at the branch, denoted as Most Valuable Clients (MVCs), did impact the market efficiency of branches.

Table 5.1
Principal component analysis summary.

Model	Significant variables	Insignificant variables[a]
Lost accounts/ Gained Accounts	Number of competitors in a10 km radius Local household income index Local population index	Number of competitors in a 1 km radius Number of competitors in a 5 km radius
Market model	Number of competitors in a 5 km radius Local household income index Local population index Number of clients holding a large amount of funds at the branch (MVCs)	Number of competitors in a 1 km radius Number of competitors in a10 km radius Number of ABMs Number of clients with large externally held funds
Delta model	Number of competitors in a10 km radius Number of competitors in a 5 km radius[b] Local household income index	Number of competitors in a 1 km radius Local population index

[a] All insignificant variables were removed from the models prior to obtaining the results provided in subsequent sections

[b] Both the 10 km and 5 km statistic were found to be significant. The two variables have a high correlation of 0.91 and when the efficiency distributions obtained from using one variable or the other were compared using a Wilcoxon Rank Sum text a ρ-statistic of 0.717 was obtained. This indicates that there is not a significant difference between the two distributions. Consequently the "Number of Competitors in a 10 km radius" variable was arbitrarily chosen and used for the remainder of the analysis.

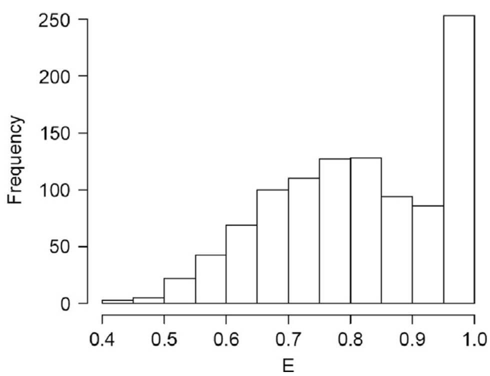

Fig. 5.1. Lost accounts/Gained accounts efficiency distribution.

5.2. Lost accounts/Gained accounts model

5.2.1. Global analysis

Of the 1040 DMUs compared in the Lost Accounts/Gained Accounts Model, 198 or approximately 19% of units were identified as efficient, with 62 or ~6% of units being identified as Pareto efficient. Although these figures are slightly lower than the general 25–30% efficient units identified in literature, this model adequately discriminated best performers from inefficient units. The efficiency distribution is provided in Fig. 5.1.

To verify the model, the growth rates of each efficient unit were calculated and compared to like units and overall averages. From this it was concluded that the Lost Accounts/Gained Accounts model provided good discriminatory power and was able to clearly identify units that outperformed others in the areas of customer retention, and new customer gains. Moreover, this model was able to calculate realistically obtainable target objectives for each branch and identified relevant peers. As is shown, in Table 5.2 the average DMU simply needs to cut employees by 2, minimize the attrition of clients by 10, the attrition of products by 17 and the attrition of funds by approximately $300,000 to operate at the efficiency level of their efficient peers. While this does not seem much, if we consider that there are over 1,000 branches, the gain could be very substantial. Although successfully attaining all of these objectives simultaneously may not be possible for some branches, they are all within a feasible range of operating goals.

Table 5.2
Lost accounts/Gained accounts average target objectives.

Input variable	Average reduction objective
Number of employees	2
Attrited customers	10
Attrited products	17
Attrited funds	$302,088

5.2.2. Local analysis

The local analyses resulted in the identification of a much larger number of efficient units. This is a result of the removal of non-comparable units that may have influenced the efficient frontier during the global analysis. These analyses were able to produce more tailored target setting objectives than the global analysis, and thus would be more useful in real world applications. The results of the local analyses are summarized in Table 5.3.

Table 5.3
Local lost accounts/Gained accounts average target objectives.

Variable	Local 1	Local 2	Local 3	Local 4	Local 5
Number of employees	1.6	2.2	0.6	2.6	1.9
Attrited customers	17.0	5.7	1.6	6.8	8.3
Attrited products	16.2	12.4	11.6	6.3	21.8
Attrited funds	211,371	211,620	17,916	403,497	309,421

When compared to each cluster's characteristics defined in Table 4.6, most of these target objectives are fitting to the surrounding demographics of the branches. For example, Cluster 3 is comprised of branches located in lower income and lower population areas which is directly reflected in the targets provided by the model.

The Lost Accounts/Gained Accounts model offers a unique perspective that provides useful insight into which branches are better at retaining customers and attracting new clients. Although the global analysis was functional, local analyses better reflect the local demographics and produce more tailored target objectives

5.3. Market models

The Market models offered the opportunity to assess the Bank branches' relative market efficiency given their unique market conditions. The Aggregate model resulted in an average efficiency score of 0.73 and 105 efficient units, 21 of which were pareto efficient. The Component model produced an average efficiency score of 0.82 and 170 efficient units, 76 of which were pareto efficient. Excluding the efficient units, the efficiency distributions of both models (Figs. 5.2 and 5.3) closely resembled the normal distribution.

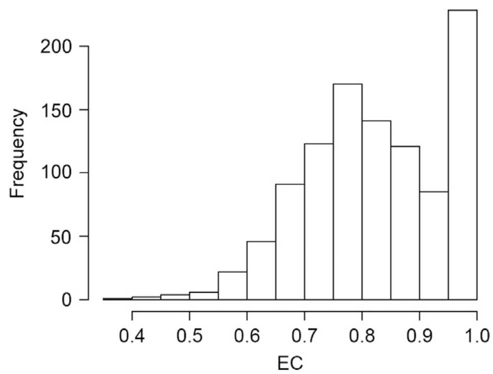

Fig. 5.2. Component market model efficiency distribution.

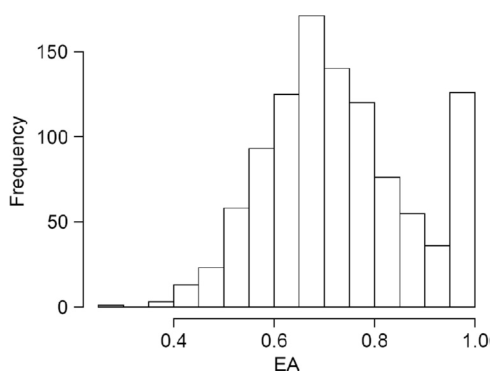

Fig. 5.3. Delta model efficiency distribution.

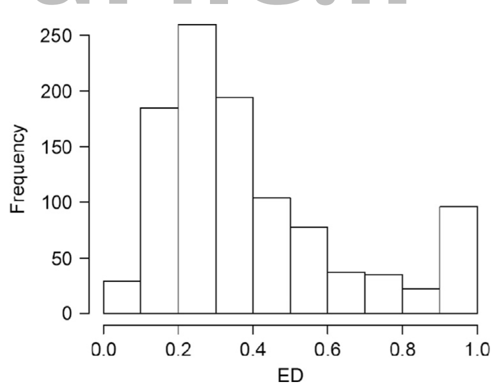

Fig. 5.4. Aggregate market model efficiency distribution.

In analyzing the results of each model, it was determined that the aggregate model consistently produced realistic and attainable target setting objectives suggesting an average increase of funds managed of approximately $7,000 per branch. The Component model appeared to be much more susceptible to the presence of speciality branches. The largest discrepancies were observed in the investment slacks, whose averages were consistently in the millions. In order to ameliorate this issue branches who deal with a high number of investments should be removed from the global analysis. For more useful local analysis results, it is probable that DMUs should be clustered into several small segments based on their distribution of product sales.

The large discrepancies that exist between the Component and Aggregate model tend to suggest that as a whole each branch holds an amount of funds that is relatively proportional to its resources; however the distribution of products from one branch to the next varies greatly. It follows that an interesting extension to these models would be to investigate product mix efficiencies and identify most efficient product mixes based on branch type and local characteristics. For this analysis to be successful the most influential branch characteristics would need to be identified and very careful segmentation would be required. Subsequently, several DEA models would need to be run to maximize efficiency. Careful attention would need to be given to slacks and target objectives to ensure they remain within a feasible range. The use of cone ratios and weight restrictions may be required to obtain viable results.

As it stands, the aggregate model was able to very clearly identify best performers and provide realistic and applicable target objectives and peer groups. The component model, however, requires further product specific segmentation in order to provide feasible results. Although this model was not initially as constructive as hoped, it sheds light on the extreme differences that exist between branch product mixes, regardless of the branches local demographics.

Table 5.4
Delta model summary statistics.

Minimum	1st Quartile	Median	Mean	3rd Quartile	Max
0.0453	0.2194	0.3216	0.40	0.5022	1.00

Table 5.5
Local delta model average output slacks.

Cluster	Difference in customers	Growth in existing products	Growth in existing funds	Difference in product count
Global	237	175	$5,477,328	877
Local 1	422	334	$16,850,910	1758
Local 2	46	377	$1,157,329	575
Local 3	285	371	$1,466,796	1072
Local 4	266	326	$3,189,664	918
Local 5	203	427	$1,187,272	950

5.4. Delta model

In order to investigate the functionally of the Delta model two model formulations were used, namely the Difference and Percent models. Despite their differences, these models produced the same efficiency scores for each DMU and thus resulted in the same efficiency distribution (Fig. 5.4).

These models identified only 70 units of 1040 as efficient and appeared much more susceptible to outliers. A summary of the model statistics are provided in Table 5.4. The low number of efficient units can most likely be attributed to the fact that each branch must operate with a pre-defined minimum number of employees. Thus many of the branches, regardless of their local market conditions and earning potential, are equipped with this base number of employees. This becomes even more apparent when local analyses are performed. Average employee reduction target objectives range from 0 to 0.8. A summary of the average output slacks for each cluster is provided in Table 5.5. On the other hand, output slacks are high, and largely lie outside of the attainable yearly growth.

It may be beneficial to use an output oriented model, although careful consideration of negative output values would be required. Moreover, information pertaining to new branch openings and relocations may aid in identifying DMUs that are experiencing higher than average growth within their first few years of operation and consequently should be removed from the PPS. Although useful in identifying branches who have realized a significant growth from one year to the next, it is overly sensitive to outliers resulting in large slack values and little constructive information in the way of target objectives. Consequently, this model should only be employed if it can be said with certainty that specialty branches and those susceptible to abnormal growth rates are not included within the PPS.

6. Conclusion

In this study, DEA methodologies were successfully applied to bank branch data to investigate three unique perspectives of branch growth. In total five models were developed using VRS DEA models to investigate which would be most successful at accurately identifying best performing branches. Through this application, it was found that two of the models successfully identified best performers and provided feasible target setting objectives. Used in conjunction with each other the Lost Account/Gained Accounts model and the Aggregate Market model can provide invaluable insight into branch customer/product churn and market efficiency. Moreover, the differences between the component and aggregate market models made it apparent that branches are highly variable in their product mixes offering the opportunity to identify most efficient product mixes given a branch's characteristics. Additionally, the use of PCA analysis provided invaluable insight into the impact of local competition, household income and population on branch efficiencies.

To continue this work, it would be interesting to evaluate whether a branch's environmental characteristics impact the efficiency of branches in rural, suburban and metropolitan regions differently. Moreover, the investigation of most efficient product mixes could be highly useful in providing more tailored marketing strategies and banking experiences. It would also be useful to investigate a means of combining the efficiency scores obtained from the developed models to obtain an overall growth efficiency for each branch. This would provide a more comprehensive indicator of branch efficiency for bank management. Moreover, the methodologies contained herein can be extended and applied to other service sectors aside from the financial industry given the availability of sufficient customer and product data over the course of several periods.

Appendix A. Raw Data Summary Statistics

See Appendix Table A1–A3

Table A1
Lost accounts/Gained accounts.

	Average	Minimum	Maximum
Input variables			
Attrited product count	420	13	2072
Attrited customers	298	12	1479
Attrited funds managed	$1644368	$5411	$15746059
Output variables			
Retained customers	5583	329	35872
Retained customer growth in product count	534	0	2334
New customers	389	17	2037
New funds managed	$3716636	$26560	$9038199

Table A2
Market model.

	Average	Minimum	Maximum
Input variables			
# of customers	5839	342	28796
Total # of most valuable customers	403	16	2101
Number of employees	12	2	80
Output variables			
Component-			
Total lending balance	$57906621	$1703987	$313364322
Total investment balance	$10982126	$6186	$64696188
Total RRSPs	$7532075	$3464	$44693186
Total RRIFs	$1252177	$1268	$9101759
Total deposits	$19363074	$64150	$127620234
Aggregate-			
Product count	18386	705	88530
Total funds managed	$96542307	$3313957	$459163963

Table A3
Delta model.

	Average	Minimum	Maximum
Input variables			
Number of employees	12	2	80
Output variables			
ΔCustomers	431	0	1338
ΔProduct count	1126	0	5029
ΔFunds managed	$11905339	$0	$72217570

*Please note the delta values have been transposed to avoid issues with negative values

References

[1] Al-Tamimi HAH, Lootah AM. Evaluating the operational and profitability efficiency of a UAE-based commercial bank. J Financ Serv Mark 2007;11:333–48.
[2] Asmild M, Tam F. Estimating global frontier shifts and global Malmquist Indices. J Prod Anal 2007;27:137–48.
[3] Athanassopoulos AD. The use of data envelopment analysis in banking institutions: evidence from the commercial bank of greece. Interfaces 2000;30:81–95.
[4] Athanassopoulos AD. Performance improvement decision aid systems in retail organizations using data envelopment analysis. J Prod Anal 1995;6:153–70.
[5] Athanassopoulos AD. Service quality and operating efficiency synergies for management control in the provision of financial services: evidence from Greek bank branches. Eur J Oper Res 1997;98:300–13.
[6] Athanassopoulos AD. Nonparametric frontier models for assessing the market and cost efficiency of large-scale bank branch networks. J. Money Credit Bank 1998;30:172–92.
[7] Banker RD, Charnes A, Cooper WW. Some models for estimating technical and scale inefficiencies in data envelopment analysis. Manage Sci 1984;30(9):1078–92.
[8] Banker RD, Morey RC. Efficiency analysis for exogenously fixed inputs and ouputes. Oper Res 1986;34(4):513–21.
[9] Berger AN. International comparisons of banking efficiency. Financ Market Institutions & Instru 2007;16:119–44.
[10] Berger AN, Humphrey DB. Efficiency of financial institutions: international survey and directions for future research. Eur J Oper Res 1997;98:175–212.
[11] Camanho AS, Dyson RG. Data envelopment analysis and malmquist indices for measuring group performance. J Prod Anal 2006;26:35–49.
[12] Camanho AS, Dyson RG. Cost efficiency, production and value-added models in the analysis of bank branch performance. J Oper Res Soc 2005;56:483–94.
[13] Canadian Bankers Association. Database of domestic banks' financial results. *Canadian Bankers Association*, 2013.
[14] Charnes A, Cooper W, Rhodes E. Measuring the efficiency of decision making units. Eur J Oper Res 1978;2(6):429–44.
[15] Colwell RJ, Davis EP. Output and productivity in banking. Scand J Econ 1992;94:111–29.
[16] Data dnvelopment analysis: a comprehensive text with models, applications, references and DEA-solver software. In: Cooper WW, Seiford LM, Tone K, editors. 2nd ed.. New York: Springer Science+Business Media; 2007.
[17] Das A, Ray SC, Nag A. Labor-use efficiency in indian banking: a branch-level analysis. Omega 2009;37:411–25.

[18] Drake L, Howcroft H. Relative efficiency in the branch network of a UK bank: an empiricalk study. Omega 1994;22:83–90.

[19] Epstein M, Henderson J. Data envelopment analysis for managerial control and diagnosis. Decis Sci 1989;20:90–119.

[20] Farrel M. The measurement of productive efficiency. J Roy Stat Soc 1957;120:253–90.

[21] Fethi MD, Pasiouras F. Assessing bank efficiency and performance with operational research and artificial intelligence techniques: a survey. Eur J Oper Res 2010;204:189–98.

[22] Gaganis C, Liadaki A, Doumpos M, Zopounidis C. Estimating and analyzing the productivity of bank branches. Managerial Financ 2009;35:202–18.

[23] Giokas DI. Bank branch operation efficiency: a comparative application of DEA and the loglinear model. Omega 1991;19:549–57.

[24] Giokas DI. Assessing the efficiency in operations of a larger Greek bank Branch network adopting different economic behaviours. Econ Model 2008;25:559–74.

[25] Golany B, Roll Y. An application procedure for DEA. Omega 1989;17(3):237–50.

[26] Kinsella R. The measurement of bank output. J I Bankers Ireland 1980;82:173–83.

[27] Manandhar R, Tang JCS. The evaluation of bank branch performance using data envelopment analysis: a framework. J High Technol Manage Res 2002;13:1–17.

[28] McEachern D, Paradi JC. Intra- and inter-country bank branch assessment using DEA. J Prod Anal 2007;27:123–36.

[29] Office of the Superintendent of Financial Institutions Canada. Fact sheet: online document, ⟨http://www.osfi-bsif.gc.ca/osfi/⟩; Cited: February 23rd, 2013.

[30] Paradi JC, Vela S, Zhu H. Adjusting for cultural differences, a new DEA model applied to a merged bank. J Prod Anal 2010;33:109–23.

[31] Paradi JC, Rouatt S, Zhu H. Two-stage evaluation of bank branch efficiency using data envelopment analysis. Omega 2011;39:99–109.

[32] Paradi JC, Zhu H. A survey on bank branch efficiency and performance research with DEA. Omega 2013;41:61–79.

[33] Parkan C. Measuring the efficiency of service operations: an application to bank branches. Eng Cost Prod Econ 1987;12:237–42.

[34] Portela MCAS, Thanassoulis E. Comparative efficiency analysis of Portuguese bank branches. Eur J Oper Res 2007;177:1275–88.

[35] Schaffnit C, Rosen D, Paradi JC. Best practice analysis of bank branches: an application of DEA in a large Canadian bank. Eur J Oper Res 1997;98:269–89.

[36] Sherman HD, Gold F. Bank branch operating efficiency: evaluation with data envelopment analysis. J Bank Financ 1985;9:297–315.

[37] Tsolas I. Modeling bank branch profitability and effectiveness by means of DEA. Int. J Prod Perform Manage 2010;59(5):432–51.

[38] Wu D, Yang Z, Liang L. Efficiency analysis of cross-region bank branches using fuzzy data envelopment analysis. Appl Math Comput 2006;181:271–81.