Cluster-Adjusted DEA Efficiency in the presence of Heterogeneity: An Application to Banking Sector

Abstract: This paper improves on the issues of extreme data points and heterogeneity found in the linear programming data envelopment analysis (DEA) by presenting a cluster-adjusted DEA model (DEA with cluster approach). This analysis, based on efficiency, determines the number of clusters via Gap statistic and Elbow methods. We use the December quarterly panel data consisting of 122 U.S agricultural banks across 37 states from 2000 to 2017 to estimate the cluster-adjusted DEA model. Empirical results show differences in the estimated DEA efficiency measures with and without a clustering approach. Furthermore, using non-parametric tests, the results of Ansari-Bradley, Kruskal Wallis, and Wilcoxon Rank Sum tests suggest that the cluster-adjusted DEA model provides statistically better efficiency measures in comparison to the DEA model without a clustering approach.

Keywords: Banking; Cluster analysis; Efficiency Analysis; Nonparametric tests

JEL: A10; C10; C14'; C44; G21

1 Introduction

Data Envelopment Analysis (DEA) is a linear programming approach that estimates a theoretical efficiency frontier based on the envelop of all the decision-making units (DMUs). Assuming monotonicity and convexity, DEA estimates the efficiency measures under alternative technology without predefined functional forms or distributional assumptions. Since its introduction by Farrell (1957), several extensions of DEA models have been proposed including: CCR model¹ (Charnes et al., 1978); BCC model² (Banker et al., 1984); Additive models (Charnes et al., 1985); Fuzzy models (Sengupta, 1992); Super efficient models (Li et al., 2007); Robust models (Shokouhi et al., 2010); Panel models (Shaik, 2013 and 2015); Quantile models (Atwood and Shaik, 2018); and Clusters or Heterogeneity models (Samoiленко and Osei-Bryson, 2008; Meiman et al., 2002; Po et al., 2009; Paradi et al., 2012; Saati et al., 2013; and Sakouvogui, 2020) with various applications in performance evaluation and benchmarking of health care (hospitals, doctors), education (schools, universities), banking sectors

---

¹CCR model is built on the assumption that regardless of operation scale an increase in inputs will result in proportional increase in outputs.
²BCC model allows more flexibility for the DEA formulation. The BCC model is formulated with the addition of the economies of scale.
(agricultural and non agricultural), manufacturing, management evaluation, energy, fast food restaurants, and retail stores (Charnes et al., 1978; Charnes et al., 1994; Wu et al., 2006; Banker and Chang, 2006; and Emrouznejad et al., 2008 and Sakouvogui, 2020).

In the presence of extreme data points or wide-spread DMUs and heterogeneity, the DEA model provides misleading and inaccurate results (Po et al., 2009). Using the statistical cluster analysis, this issue of extreme data points and heterogeneity has been addressed (Samoilenko and Osei-Bryson, 2008; Jessop, 2009; and Po et al., 2009). However, defining clusters based on an input-output frontier provides biased and higher efficiency measures due to the localized frontier defined by the input-output clusters (Meiman et al., 2002 and Saatii et al., 2013). Following Shaik et al., (2012), this bias is due to three reasons: 1) The input-output frontier is driven by extreme DMUs. These DMUs force the efficiency to be 1 or closer to 1 suggesting these are the most efficient DMUs; 2) The weights associated with extreme DMUs force them to be 1; and 3) In order to identify clusters based on input-output frontier would require to identify clusters based on output or each individual inputs. Thus, the estimated efficiency accounts for potential clustering in output and inputs, so advantageous in clustering the data using efficiency rather than output-inputs or the associated frontier.

In this paper, we present one possibility to improve the DEA model by developing a cluster-adjusted DEA estimator based on a four-step methodology. The four-steps involve estimation of efficiency measures, cluster analysis based on efficiency, re-estimation of efficiency measures by cluster, and nonparametric statistical tests. First, using the DEA model without a clustering approach, we estimate the efficiency measures of the DMUs by year. Second, based on the estimated efficiency measures, we determine the optimal number of cluster groups. Third, using the statistically identified clusters of DMUs, we estimate the cluster-adjusted DEA model by year. Finally, we provide nonparametric tests to evaluate differences in the DEA efficiency measures estimated with and without a clustering approach using Kolmogorov-Smirnov Statistics, Ansari-Bradley, Kruskal Wallis, and Wilcoxon Rank Sum tests.

Following the introduction in Section 1, the rest of this paper is organized into five sections. Section 2 presents a brief literature review. Section 3 discusses the theoretical framework of the DEA and cluster-adjusted DEA models. Section 4 presents the empirical data and construction of the input and output variables. Section 5 discusses the results and statistical implications. Finally, the summary of our conclusions is presented in Section 6.

2 Literature Review

Since the seminal works of Koopmans (1951), Farell (1957), and Farrell and Fieldhouse (1962) on efficiency measures, DEA of Charnes et al., (1978), a non-parametric method used for the estimation of production frontiers, can evaluate the performance of DMUs by measuring the relative efficiency using multiple inputs and multiple outputs. This measurement can be either output-oriented or input-oriented. With its application in the banking sector (Aly et al., 1990; Miller and Noulas, 1996; Chen 2002; Casu and Molyneux, 2003; Drake and Hall, 2003; Hauner, 2005; Tao et al., 2013; and Dipayan, 2014), DEA has expanded in dealing with the issues of heterogeneity in DMUs by integrating the DEA model and cluster analysis to alleviate the gaps in DEA modeling (Yang and Kuo, 2003; Paradi et al., 2012; and Maletic et al., 2013).

For example, Thanassoulis (1996) developed a method for simultaneously clustering operating units and determining a different set of marginal resource levels. Meiman et al., (2002) applied clustering directly to the results of DEA method with the goal of having multiple references of subsets from the original set of DMUs. Po et al., (2009) employed a piecewise production functions derived from the DEA method to cluster the data with input and output variables. Jessop (2009) used an integer DEA model with both the number and size distribution of groups as objectives and criteria. Samoilenko and Osei-Bryson (2008) presented a three steps methodology: cluster analysis, DEA, and Decision tree.

Alirezaee and Sani (2011) presented a new hierarchical process DEA model. Paradi et al., (2012) applied the k-means clustering algorithm with DEA efficiency measures. Amirteimoori and Kordrostami (2014) clustered the operational units then evaluated each unit in its cluster. Tao et al., (2013) presented a hybrid model
to conducting performance measurements using DEA and axiomatic fuzzy set clustering. Jahangoshai et al., (2018) integrated dynamic fuzzy C-means and Artificial Neural Network with a DEA model to solve a multiple criteria optimization problem.

Our paper differs from the existing literature in four different aspects. First, unlike the existing literature such as Tone (2017) that defined the clusters based on an input-output frontier, this paper addresses the issue of heterogeneity by first estimating the efficiency measures through the concept of DEA. Second, the optimal number of clusters is identified using alternative clustering methods, Gap Statistic and Elbow methods. The results of the optimal number of cluster groups are then compared to the different cluster indices discussed in Charrad et al., (2014). Third, the efficiency measures are re-estimated by cluster groups while accounting for the yearly variability. Finally, to test the robustness of the DEA model, nonparametric statistical tests are used to evaluate distributions of the efficiency measures estimated with and without the clustering method.

3 Theoretical framework

Primal production theory assumes that the relationship between multiple outputs, \( y = (y_1, x_2, \ldots, y_j) \in \mathbb{R}^j \) and inputs, \( x = (x_1, x_2, \ldots, x_i) \in \mathbb{R}^i \) is reflected by the concept of production function. The production function represents the relation between non-allocatable exogenous input vectors, \( x \), used in the production of an endogenous output, \( y \). The production function framework forms the bases in the estimation of the DMU’s efficiency using the linear programming DEA model.

3.1 DEA model

The technology that transforms multiple inputs into multiple outputs is represented by input set, \( L(y) \). The input set, \( L(y) \), satisfying constant returns to scale and strong disposability of input is defined as:

\[
L(y) = \{ x : y \text{ can produce } x; \quad x \in \mathbb{R}^i \text{ and } y \in \mathbb{R}^j \} \tag{1}
\]

The input set, \( L(y) \), denotes the collection of input vectors that yield output vectors. This concept is represented by an input distance function evaluated for any DMU reference production possibility set, \( T \), as:

\[
D_T(y^t, x^t) = \min \{ \lambda : (\lambda x^t \in L^T(y^t)) \} \tag{2}
\]

or

\[
\min_{\delta z} \quad y^t \leq Yz, \quad Y = y_1, \ldots, y_T.
\]

\[
\lambda x^t \geq Xz, \quad X = x_1, \ldots, x_T.
\]

\[
z \geq 0.
\]

Here, the second expression of equation (2) identifies the linear program used to calculate the distance function, with \( z \) being a \( T \times 1 \) vector of intensity variables. Therefore, \( z \) identifies the constant returns-to-scale (CRS) boundaries of the reference set. Under the variable returns-to-scale (VRS), the intensity variable is \( z=1 \). In addition, the scale efficiency measure is computed as the ratio of the efficiency measure estimated under CRS over pure technology estimated under VRS.

The efficiency measures estimated by the DEA model (equation 2) forms the basis for a cluster analysis. The optimal number of cluster groups is identified using alternative clustering methods, Gap Statistic and Elbow methods. Next, the conceptual framework of the clustering method is presented.

3.2 Estimating the number of clusters based on DEA efficiency measures

Cluster analysis deals with the identification of homogenous DMUs with similar patterns. Suppose that we have already estimated the efficiency measures of the DEA model. Let \( \lambda_{ij} \), \( t = 1, \ldots, T \) define indepen-
dent efficiency measures and \( j = 1, \ldots, p \) DMU’s. Additionally, suppose that we have already clustered the efficiency measures \( \lambda_{ij} \) into \( k \) clusters \( C_1, \ldots, C_k \) with \( C_k \) denoting the indices of observations in \( k \) and \( n_k = |C_k| \). Thereafter, let \( d_{pp'} \) be the square Euclidean distance between the two DMUs, \( p \) and \( p' \), of efficiency measure, \( \lambda \). There are several different indices for choosing the optimal number of clusters, \( k \), in the \( k \)-means method, among them we focus on Gap Statistic and Elbow methods. By the rule of thumb, we assume that the maximum \( k \) value is: \( k_{\text{max}} = \sqrt{\frac{p}{2}} \).

### 3.2.1 Gap statistic method

The gap statistic method, first by developed by Tibshirani et al., (2001) is based on the log standardization of the pooled within cluster sum of square, \( \log W_k \). The determination of the number of cluster groups is as follows:

- **Step 1:** From the number of clusters, \( k = 1, \ldots, k_{\text{max}} \), compute the pooled within cluster sum of squares around the cluster means, \( W_k \), as:
  \[
  W_k = \sum_{k=1}^{k_{\text{max}}} \frac{D_k}{2n_k}
  \]
  where \( D_k = \sum_{p,p' \in C_k} d_{pp'} \) is the sum of the pairwise distance for all the points in cluster, \( k \) and \( n_k \) is the number of DMUs in cluster, \( k \).
- **Step 2:** Generate \( B \) reference data sets using the uniform distribution and then cluster each data set with a varying number of clusters, \( k = 1, \ldots, k_{\text{max}} \).
- **Step 3:** Calculate the within-dispersion measures, \( W'_{kb} \) with \( b = 1, \ldots, B \). Then compute the estimated Gap statistic
  \[
  \text{Gap}(k) = \left( \frac{1}{B} \right) \sum_{b=1}^{B} \log(W'_{kb}) - \log(W_k)
  \]
- **Step 4:** Let \( \bar{\nu} = \left( \frac{1}{B} \right) \sum_{b=1}^{B} \log(W'_{kb}) \) and the standard deviation, \( s_d(k) \), be defined as: \( s_d(k) = \left[ \frac{1}{B} \left( \sum_{b=1}^{B} \log(W'_{kb}) - \bar{\nu} \right)^2 \right]^{1/2} \) and define \( s_k = \sqrt{1 + \frac{1}{B} s_d(k)} \).
- **Step 5:** Choose the number of clusters as the smallest \( k \) such as \( \text{Gap}(k) \geq \text{Gap}(k) - s_{(k+1)} \).

### 3.2.2 Elbow method

One the oldest methods for determining the optimal number of clusters is the Elbow method (Sugar, 1998). The algorithm can be computed as follows:

- **Step 1:** With a varying number of clusters, \( k = 1, \ldots, k_{\text{max}} \), compute the clustering algorithm using the DEA efficiency scores.
- **Step 2:** For each \( k \), compute the pooled within cluster sum of squares around the cluster means, \( W_k \), as:
  \[
  W_k = \sum_{k=1}^{k_{\text{max}}} \frac{D_k}{2n_k}
  \]
  where \( D_k = \sum_{p,p' \in C_k} d_{pp'} \) is the sum of the pairwise distance for all the points in cluster, \( k \) and \( n_k \) is the number of DMUs in cluster, \( k \).
- **Step 3:** At some value of \( k \), \( W_k \) will drop dramatically. Thereafter, it will reach a diminishing return with an increase in \( k \). Therefore, choose \( k \) that does not increase much \( W_k \).

---

3 Additionally, to provide a more robust optimal number of clusters, we additionally provide the results of the majority votes for the 30 different indices in the R package NbClust and show that Gap Statistic and Elbow indices are sufficient.
The optimal number of clusters, $k$, identified by Gap Statistic and Elbow methods forms the basis for the cluster-adjusted DEA model. These techniques use the efficiency measures of the DEA model and the output of the k-means clustering algorithm to form homogenous DMUs. The efficiency of these clusters depends on the change of the within-cluster dispersion, $W_k$. Hence, taking into consideration the predefined minimum and maximum $k$ values in determining the optimal $k$ value around the DMUs, the cluster of efficiency measures will result in dynamic size of homogenous DMUs.

### 3.3 Cluster-adjusted DEA model

Once the optimal $k$ value and number of homogenous DMUs are identified based on the efficiency measures, the cluster-adjusted DEA model is estimated using equation 3. The technology that transforms inputs into outputs is represented by the cluster-input set, $L(y^k)$, where $k$ is the number of cluster groups. Thus, the cluster-input distance function evaluated for any DMU within each cluster-reference production possibility set, $K$, is expressed as:

$$D^k_i(y^k, x^k)^{-1} = \min\{\lambda : \lambda x^k \in L^k(y^k)\}$$

(3)

or

$$\min_{\theta z} \quad \text{subject to} \quad y^k \leq Yz, \; Y = y_1, \ldots, y_K, \quad \lambda x^k \geq Xz, \; X = x_1, \ldots, x_K, \quad z \geq 0,$$

where the number of clusters, $k$, and all other properties of the input distance function remains the same.

The empirical application of our method is straightforward. The estimation of the DEA estimator in an efficiency-cluster-efficiency set up is completed as follows:

1. Estimate the input-oriented DEA model by year (equation 2) under CRS, VRS, and scale assumptions.
2. Determine the optimal number of cluster groups by year.
3. Cluster the DEA efficiency measures while accounting for the yearly variability (technological change).
4. Estimate the cluster-adjusted DEA model by year (equation 3) under CRS, VRS, and scale assumptions.
5. Conduct nonparametric tests to evaluate differences in DEA efficiency measures estimated with and without a clustering approach.

### 4 Empirical data

The agricultural banking sector is a major component of the financial system. Hence, performance of the agricultural banking sector is critical to the stability and development of the United States' (U.S) economy. One of the most important factors that affect the agricultural banking performance is the interest rates. An analysis of bank interest rate determinants is crucial to the understanding of the financial intermediation. Interest rate is the price a borrower pays for the use of money they borrow from a lender or fee paid on borrowed assets (Crowley, 2007). Interest rates determine the profitability of agricultural banks among other factors (Gardner et al., 2005).

Proper interest rate management reduces bank exposure to risk and provides an opportunity to stabilize and improve their net income. According to Flannery (1980), when interest rates rise, banks managers can

---

4 The gap is part of the NbClust package which consists of 30 different indexes used to determine the optimal number of clusters. Hence, the determination of the optimal cluster is carried through the selection of the winner via the majority voting scheme of NbClust package. Nbclust is available on:https://cran.r-project.org/web/packages/NbClust/NbClust.pdf.
expect changes on both the asset and liability sides of their balance sheets. Since banks are always faced with the risk of having a low or high interest rate, the application of our methodology to the agricultural banking sector could help managers and regulators make sound banking decisions. Furthermore, in analyzing the interest rates of agricultural banks, we are addressing a significant issue in banking - banks are more sensitive to the changes in interest rates.

The Farm Credit Administration (FCA), an institutional part of the United States government provides a uniform call report that contains the financial data of each Farm Credit System or agricultural banks that must be submitted to FCA quarterly. The data is freely available at: https://www.fca.gov/bank-oversight/call-report-data-for-download. After a post-hoc data cleaning, this paper uses a December quarterly data consisting of 122 agricultural banks across 37 states from 2000 to 2017. The input and output variables obtained from FCA must be consistent and provide a true reflection of the interest rates.

The output variables were selected to represent the income side of the agricultural banking sector. Thus, the selected two output variables were defined as the total interest income and the total non-interest income. The input variables included were the total interest expenses and the total non-interest expenses.

The selection of input and output variables in the DEA model needs careful attention because it may affect the distribution of technical efficiency measures. Since income is output based, the output price index is used to deflate the income. Thus, the deposit service (DS) and loan service (LS) price indexes are used to compute the quantity index (QI) of the output and input variables, respectively. The quantity index of the output and input variables is computed as, QI (output)= aggregate output variable × (100/LS) and QI (input)= input variable × (100/DS), respectively. The aggregate output variable is defined as a sum of total interest income and total noninterest income. Table 1 presents the summary statistics of inputs, outputs, aggregate output, and price indexes. Since, FCA reports data that contains negative input values, the application of DEA model cautions the researcher to first convert the negative values into positive by adding a common positive number. Thus, with a negative interest expense, a constant of 9,500 was added to the quantity indices (Table 1).

---

5 The quarterly data for December would include the previous 3 quarters data. The physical year starts with January 1 and ends with December 31, of each year.

6 The total interest income (RIAD4017) is the sum of 1) Total interest and fee income on loans (RIAD4010); 2) Income from lease financing receivables (RIAD4065); 3) Interest income on balances due from depository institutions (RIAD4115); 4) Interest and dividend income on: U.S. Treasury securities and U.S. Government agency obligations excluding mortgage-backed securities (RIADB488), Mortgage-backed securities (RIADB489), and all other securities includes securities issued by states and political subdivisions (RIADB460); 5) Interest income from trading assets (RIAD4069); 6) Interest income on federal funds sold and securities purchased under agreements to resell (RIAD4020); and 7) Other interest income (RIAD4518). The total non-interest income (RIAD4079) is the sum of 1) Income from fiduciary activities (RIAD4070); 2) Service charges on deposit accounts (RIAD4080); 3) Trading revenue (RIADA220); 4) Fees and commissions from securities brokerage (RIADC886); 5) Investment banking, advisory, and underwriting fees and commissions (RIADC888); 6) Fees and commissions from annuity sales (RIADC887); 7) Underwriting income from insurance and reinsurance activities (C386); 8) Income from other insurance activities (RIADC387); 9) Venture capital revenue (RIADB491); 10) Net servicing fees (RIADB492); 11) Net securitization income (RIADB493); 12) Net gains (losses) on sales of loans and leases (RIAD5416); 13) Net gains (losses) on sales of other real estate owned (RIAD5415); 14) Net gains (losses) on sales of other assets (RIADB496); and 15) Other noninterest income (RIADB497).

7 The total interest expense (RIAD4073) is the sum of 1) Interest on deposits in domestic offices on: Transaction accounts (interest-bearing demand deposits, NOW accounts, ATS accounts, and telephone and preauthorized transfer accounts (RIAD4508); Non-transactions accounts with savings deposits (includes MMDA’s) (RIADB093), Non-transactions accounts with time deposits of $250,000 or less (RIADBHK03), Non-transactions accounts with time deposits of more than $250,000 (RIADBHK04); 2) Interest on deposits in foreign offices, edge and agreement subsidiaries, and IBFs (RIAD4172); 3) Expense of federal funds purchased and securities sold under agreements to repurchase (RIAD4180); 4) Interest on trading liabilities and other borrowed money (RIAD4185); and 5) Interest on subordinated notes and debentures (RIAD4200). The total non-interest expense (RIAD4093) is the sum of Non-interest expense on: 1) Salaries and employee benefits (RIAD4135); 2) Expenses of premises and fixed assets (net of rental income) (excluding salaries and employee benefits and mortgage interest (RIAD4217); 3) Goodwill impairment losses (RIADC216); 4) Amortization expense and impairment losses for other intangible assets (RIADC232); and 5) Other noninterest expense (RIAD4092).
Table 1: Descriptive Statistics of the input and output variables in indexes.

| Variable            | Mean   | Standard deviation | Maximum   | Minimum  |
|---------------------|--------|--------------------|-----------|----------|
| **Quantity index**  |        |                    |           |          |
| Total interest expenses | 22,615 | 60,765.31          | 655,077   | 58       |
| Total noninterest expenses | 9,601  | 16,842.91          | 144,968   | -9,461   |
| Total interest income | 26,206.40 | 59,076.97        | 667,613.60 | 218.2   |
| Total noninterest income | 4,303.64 | 7,116.89          | 70,873.31 | 3.07     |
| Aggregate Output    | 30,510.10 | 62,967.22          | 677,599.60 | 221.2   |
| **Index Measures**  |        |                    |           |          |
| Loans services      | 123.35 | 23.02              | 160.5     | 91.83    |
| Deposits services   | 82.18  | 21.15              | 118.7     | 56.3     |

The input and output variables are in thousands of dollars.

5 Empirical Results and Discussions

An input-oriented DEA model was adopted because 1) Banks have better control over inputs than outputs and 2) In the presence of negative input values, the output-oriented DEA model becomes infeasible. Using the quantity index of inputs and the aggregate output variables, the DEA model (equation 2) was estimated by year due to the differences in banks through time under CRS, VRS, and scale assumptions. While accounting for the yearly variability, the optimal number of clusters and the partition of banks into groups were done using the estimated DEA efficiency measures under CRS assumption. In addition, the cluster-adjusted DEA model was estimated using equation 3. All the models were estimated in R language and the nonparametric tests in Statistical Analysis Software.

5.1 Efficiency measures

The efficiency measures estimated using the DEA model defined in equation (2) are used in the cluster analysis. Table 2 presents the summary statistics of the DEA efficiency measures estimated under CRS and VRS assumptions. The DEA efficiency measures under the scale assumption were estimated as the ratio of DEA efficiency measures under CRS to DEA efficiency measures under VRS assumptions. Three important results emerge from Table 2.

First, the yearly mean DEA efficiency measures range from 0.863 to 0.900 under CRS, 0.933 to 0.953 under VRS and 0.920 to 0.952 under the scale assumptions. The results in mean DEA efficiency measures are validated by the slight fluctuation of the standard deviations through the years. Furthermore, the results suggest that during the financial crisis of 2007-2009, banks were on average efficient. Second, the limitation of a standard formulation of the DEA model is to build a separate linear program for each bank. However, since the data is composed of wide-spread and heterogenous banks, the efficiency measures estimated in equation 2 are biased and inaccurate. That is, the DEA model without a clustering approach fails to define the group of banks that are like the banks under evaluation. Henceforth, it may be difficult to interpret the results of Table 2 because of the non-homogeneous banks. Finally, a comparison of the DEA efficiency measures under the scale assumption suggests that the VRS technology is higher than CRS technology. To avoid bias due to scale efficiency, the optimal number of clusters is determined based on the CRS efficiency measures.
package suggest that four cluster groups are sufficient to partition the banks. Second, even though the num-
groups are sufficient to partition the efficiency measures of banks. These results are further validated by the
Elbow methods, and the 30 indices present in the NbClust package. Within each year of Table 3, three cluster
variability of the efficiency measures. Hence, with distinct groups of efficiency measures presented annually
output variables, the results in Figure 1 illustrate the presence of cluster groups and the importance of yearly
PD

Existing literature has shown that the analysis of banks’ performance and efficiency would be almost
impossible if all banks had the same capital structure, offered the same mix of services, followed identical
accounting practices, and were equally affected by inflation and operated under the same regulatory restric-
tions (Vittas, 1991). However, with banks exhibiting considerable differences in their efficiency measures,
qualitative problems would undermine the usefulness of these ratios for analytical and policy purposes when
not properly accounted for.

Figure 1 presents the yearly distribution by year of the DEA efficiency measures under CRS assumption.
Unlike the existing literature such as Meiman et al., (2002) and Saati et al., (2013) who clustered the input and
output variables, the results in Figure 1 illustrate the presence of cluster groups and the importance of yearly
variability of the efficiency measures. Hence, with distinct groups of efficiency measures presented annually
in Figure 1, Table 3 summarizes the results of the k-means clustering approach based on the Gap statistic and
Elbow methods, and the 30 indices present in the NbClust package. Within each year of Table 3, three cluster
groups are sufficient to partition the efficiency measures of banks. These results are further validated by the
majority rule decision of the 30 indices. In addition, from Table 3, two important results emerge.

First, 12 indices of the NbClust package suggest that within each year, the optimal number of clusters is
three. This result is based on the majority vote or ruling of the NbClust. In addition, 5 indices of the NbClust
package suggest that four cluster groups are sufficient to partition the banks. Second, even though the num-

Table 2: Summary Statistics of the efficiency measures.

| Year | Number of banks | DEA efficiency measures under CRS assumption | DEA efficiency measures under VRS assumption | DEA efficiency measures under scale assumption |
|------|-----------------|---------------------------------------------|---------------------------------------------|---------------------------------------------|
|      |                 | Mean | Std.dev | Minimum | Mean | Std.dev | Minimum | Mean | Std.dev | Minimum | Mean | Std.dev | Minimum |
| 2000 | 56              | 0.887 | 0.079 | 0.671 | 0.733 | 0.067 | 0.466 | 0.667 | 0.09 | 0.09 | 0.624 | 0.94 | 0.066 | 0.667 |
| 2001 | 58              | 0.9     | 0.093 | 0.627 | 0.733 | 0.062 | 0.466 | 0.733 | 0.073 | 0.073 | 0.533 | 0.936 | 0.066 | 0.533 |
| 2002 | 58              | 0.873     | 0.113 | 0.351 | 0.668 | 0.087 | 0.466 | 0.729 | 0.054 | 0.054 | 0.781 | 0.95 | 0.068 | 0.781 |
| 2003 | 60              | 0.863     | 0.1 | 0.502 | 0.807 | 0.069 | 0.466 | 0.814 | 0.068 | 0.068 | 0.791 | 0.947 | 0.068 | 0.791 |
| 2004 | 60              | 0.876     | 0.089 | 0.673 | 0.746 | 0.069 | 0.466 | 0.879 | 0.066 | 0.066 | 0.799 | 0.942 | 0.066 | 0.799 |
| 2005 | 60              | 0.87     | 0.087 | 0.668 | 0.668 | 0.087 | 0.466 | 0.879 | 0.068 | 0.068 | 0.596 | 0.942 | 0.068 | 0.596 |
| 2006 | 60              | 0.885     | 0.087 | 0.626 | 0.668 | 0.087 | 0.466 | 0.879 | 0.081 | 0.081 | 0.655 | 0.891 | 0.081 | 0.655 |
| 2007 | 60              | 0.877     | 0.093 | 0.648 | 0.668 | 0.087 | 0.466 | 0.879 | 0.091 | 0.091 | 0.677 | 0.887 | 0.091 | 0.677 |
| 2008 | 60              | 0.88     | 0.088 | 0.695 | 0.737 | 0.063 | 0.466 | 0.941 | 0.057 | 0.057 | 0.79 | 0.941 | 0.057 | 0.79 |

Std.dev: Standard deviation. Estimation is based on the input-oriented DEA model.
ber of cluster groups is identical from 2000 to 2017, the composition of banks within cluster differed by year (dynamic). For example, for the first cluster group, 27 banks were in 2000 whereas 22 banks in 2001. In contrast, for the second cluster group, 11 banks were in 2000 and 17 banks in 2001. After statistically determining the optimal number of clusters, the cluster-adjusted DEA model (equation 3) is re-estimated while accounting for the yearly variability. Overall, most banks experienced an increase in their efficiency measures in the early 2005 and followed a downward trend during the financial crisis of 2007. Additionally, the results of Table 2 clearly depict that most banks were able to convert inputs to outputs effectively.

Table 3: Optimal Number of clusters by year.

| Year | $k=2$ | $k=3$ | $k=4$ | $k=5$ | $k=6$ | Gap statistic | Elbow method |
|------|-------|-------|-------|-------|-------|---------------|--------------|
| 2000 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2001 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2002 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2003 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2004 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2005 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2006 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2007 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2008 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2009 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2010 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2011 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2012 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2013 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2014 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2015 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2016 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |
| 2017 | 3     | 12    | 5     | 2     | 2     | 3             | 3            |

Majority rule: Number of indices for a given $k$ value.
Best $k$ value based on Gap Statistic and Elbow methods.
Figure 1: DEA efficiency measures by year under CRS assumption.
5.2 Cluster-adjusted efficiency measures

Having first discussed the estimation of the DEA efficiency measures (Table 2), it is important to know the magnitude of the efficiency measures estimated using the cluster-adjusted DEA model. Tables 4, 5, and 6 respectively present the summary statistics of the cluster-adjusted DEA efficiency measures by year within cluster 1, 2, and 3. In comparison to Table 2, the cluster-adjusted DEA efficiency measures in Tables 4, 5, and 6 are higher while accounting for homogenous banks. The higher efficiency measures do not preclude that the banks are performing better. However, this suggests that we are evaluating homogenous banks or DMUs based on similar characteristics identified by the cluster analysis of efficiency.

**Table 4:** Summary Statistics of efficiency measures within cluster group 1.

| Year | Number of banks | Mean  | Std.dev | Minimum | Year | Number of banks | Mean  | Std.dev | Minimum |
|------|-----------------|-------|---------|---------|------|-----------------|-------|---------|---------|
|      | DEA efficiency measures under CRS assumption |       |         |         |      | DEA efficiency measures under VRS assumption |       |         |         |
| 2000 | 27              | 0.974 | 0.013   | 0.953   | 2009 | 23              | 0.858 | 0.116   | 0.631   |
| 2001 | 22              | 0.818 | 0.078   | 0.733   | 2010 | 15              | 0.895 | 0.098   | 0.651   |
| 2002 | 23              | 0.973 | 0.019   | 0.942   | 2011 | 25              | 0.934 | 0.058   | 0.845   |
| 2003 | 23              | 0.885 | 0.062   | 0.787   | 2012 | 24              | 0.923 | 0.075   | 0.783   |
| 2004 | 19              | 0.926 | 0.053   | 0.827   | 2013 | 24              | 0.979 | 0.025   | 0.92     |
| 2005 | 15              | 0.922 | 0.063   | 0.818   | 2014 | 20              | 0.882 | 0.165   | 0.352   |
| 2006 | 21              | 0.901 | 0.074   | 0.706   | 2015 | 20              | 0.903 | 0.073   | 0.809   |
| 2007 | 20              | 0.888 | 0.093   | 0.093   | 2016 | 25              | 0.898 | 0.082   | 0.668   |
| 2008 | 22              | 0.923 | 0.063   | 0.797   | 2017 | 22              | 0.945 | 0.041   | 0.804   |

| Year | Number of banks | Mean  | Std.dev | Minimum | Year | Number of banks | Mean  | Std.dev | Minimum |
|------|-----------------|-------|---------|---------|------|-----------------|-------|---------|---------|
|      | DEA efficiency measures under VRS assumption |       |         |         |      | DEA efficiency measures under scale assumption |       |         |         |
| 2000 | 27              | 0.99  | 0.01    | 0.96    | 2009 | 23              | 0.931 | 0.065   | 0.821   |
| 2001 | 22              | 0.96  | 0.044   | 0.852   | 2010 | 15              | 0.945 | 0.061   | 0.834   |
| 2002 | 23              | 0.978 | 0.018   | 0.942   | 2011 | 25              | 0.991 | 0.016   | 0.936   |
| 2003 | 23              | 0.964 | 0.037   | 0.895   | 2012 | 24              | 0.964 | 0.045   | 0.846   |
| 2004 | 19              | 0.94  | 0.052   | 0.86    | 2013 | 24              | 0.989 | 0.017   | 0.941   |
| 2005 | 15              | 0.94  | 0.048   | 0.858   | 2014 | 20              | 0.959 | 0.066   | 0.805   |
| 2006 | 21              | 0.942 | 0.057   | 0.799   | 2015 | 20              | 0.989 | 0.017   | 0.932   |
| 2007 | 20              | 0.987 | 0.022   | 0.919   | 2016 | 25              | 0.971 | 0.043   | 0.868   |
| 2008 | 22              | 0.972 | 0.041   | 0.847   | 2017 | 22              | 0.979 | 0.038   | 0.866   |

Std.dev: Standard deviation. Estimation is based on the input-oriented DEA model.
Table 5: Summary Statistics of efficiency measures within cluster group 2.

| Year | Number of banks | Mean  | Std.dev | Minimum | Year | Number of banks | Mean  | Std.dev | Minimum |
|------|-----------------|-------|---------|---------|------|-----------------|-------|---------|---------|
|      | DEA efficiency measures under CRS assumption |       |         |         |      | DEA efficiency measures under VRS assumption |       |         |         |      | DEA efficiency measures under scale assumption |       |         |         |      |
| 2000 | 11              | 0.933 | 0.061   | 0.802   | 2009 | 15              | 0.929 | 0.055   | 0.831   |      | 2000              | 11              | 0.971 | 0.035   | 0.893   | 2009 | 15              | 0.957 | 0.05   | 0.855   |
| 2001 | 17              | 0.836 | 0.083   | 0.736   | 2010 | 24              | 0.906 | 0.107   | 0.514   |      | 2001              | 17              | 0.951 | 0.042   | 0.878   | 2010 | 24              | 0.952 | 0.061   | 0.806   |
| 2002 | 12              | 0.979 | 0.02    | 0.944   | 2011 | 17              | 0.953 | 0.029   | 0.913   |      | 2002              | 12              | 0.975 | 0.037   | 0.905   | 2011 | 17              | 0.987 | 0.016   | 0.957   |
| 2003 | 16              | 0.867 | 0.058   | 0.805   | 2012 | 18              | 0.927 | 0.058   | 0.795   |      | 2003              | 16              | 0.948 | 0.052   | 0.816   | 2012 | 14              | 0.956 | 0.056   | 0.793   |
| 2004 | 19              | 0.919 | 0.067   | 0.74    | 2013 | 14              | 0.911 | 0.086   | 0.674   |      | 2004              | 19              | 0.975 | 0.027   | 0.91    | 2013 | 19              | 0.965 | 0.037   | 0.899   |
| 2005 | 27              | 0.953 | 0.037   | 0.883   | 2014 | 19              | 0.928 | 0.058   | 0.782   |      | 2005              | 27              | 0.942 | 0.057   | 0.825   | 2014 | 15              | 0.983 | 0.018   | 0.938   |
| 2006 | 20              | 0.858 | 0.1     | 0.732   | 2015 | 15              | 0.885 | 0.051   | 0.82    |      | 2006              | 20              | 0.981 | 0.019   | 0.944   | 2015 | 15              | 0.959 | 0.043   | 0.866   |
| 2007 | 19              | 0.974 | 0.02    | 0.934   | 2016 | 15              | 0.913 | 0.067   | 0.73    |      | 2007              | 19              | 0.921 | 0.092   | 0.704   | 2016 | 20              | 0.983 | 0.029   | 0.878   |
| 2008 | 22              | 0.827 | 0.108   | 0.65    | 2017 | 20              | 0.891 | 0.08    | 0.78    |      | 2008              | 22              | 0.961 | 0.061   | 0.802   | 2017 | 15              | 0.971 | 0.043   | 0.831   |

Std.dev: Standard deviation. Estimation is based on the input-oriented DEA model.
Table 6: Summary Statistics of efficiency measures within cluster group 3.

| Year | Number of banks | Mean | Std.dev | Minimum | Year | Number of banks | Mean | Std.dev | Minimum |
|------|-----------------|------|---------|---------|------|-----------------|------|---------|---------|
|      | DEA efficiency measures under CRS assumption |      |         |         |      | DEA efficiency measures under VRS assumption |      |         |         |      | DEA efficiency measures under scale assumption |
| 2000 | 18              | 0.911| 0.1     | 0.733   | 2000 | 18              | 0.977| 0.03    | 0.859   | 2000 | 18              | 0.932| 0.092   | 0.748   |
| 2001 | 19              | 0.909| 0.07    | 0.799   | 2001 | 19              | 0.981| 0.04    | 0.828   | 2001 | 19              | 0.928| 0.068   | 0.799   |
| 2002 | 23              | 0.94 | 0.049   | 0.872   | 2002 | 23              | 0.985| 0.017   | 0.946   | 2002 | 23              | 0.955| 0.044   | 0.884   |
| 2003 | 21              | 0.913| 0.075   | 0.786   | 2003 | 21              | 0.966| 0.025   | 0.924   | 2003 | 21              | 0.944| 0.069   | 0.801   |
| 2004 | 22              | 0.894| 0.071   | 0.747   | 2004 | 22              | 0.931| 0.066   | 0.762   | 2004 | 22              | 0.96 | 0.043   | 0.81    |
| 2005 | 18              | 0.938| 0.055   | 0.811   | 2005 | 18              | 0.959| 0.048   | 0.837   | 2005 | 18              | 0.977| 0.03    | 0.88    |
| 2006 | 19              | 0.808| 0.062   | 0.706   | 2006 | 19              | 0.955| 0.039   | 0.88    | 2006 | 19              | 0.955| 0.039   | 0.88    |
| 2007 | 21              | 0.965| 0.029   | 0.902   | 2007 | 21              | 0.985| 0.017   | 0.95    | 2007 | 21              | 0.985| 0.017   | 0.95    |
| 2008 | 16              | 0.833| 0.147   | 0.624   | 2008 | 16              | 0.892| 0.116   | 0.67    | 2008 | 16              | 0.931| 0.069   | 0.802   | 0.802|

Std.dev: Standard deviation. Estimation is based on the input-oriented DEA model.

5.3 Nonparametric tests of difference in efficiency with and without the clustering approach

This subsection focuses on the statistical comparison of the DEA efficiency measures estimated with and without a clustering approach. The statistical comparison involves the following: 1) compute the magnitude of differences in efficiency measures with and without a clustering approach, and 2) conduct statistical tests to evaluate the significance of differences in efficiency measures with and without a clustering approach. The first analysis, accomplished by computing the change in the efficiency measures, is defined as:

$$
\delta_{it} = \hat{\mu}_{it} - \hat{\mu}_{it},
$$

where $\hat{\mu}_{it}$ and $\hat{\mu}_{it}$ are respectively the efficiency measures estimated with and without a clustering approach, and $i$ represents the individual bank, and $t$ represents the time.

To evaluate differences in efficiency measures estimated with and without a clustering approach, we investigate the feasibility of the parametric approach to statistically test the significance level of $\delta_{it}$. Figures 2, 3, and 4 respectively show the distributions of $\delta_{it}$ for the CRS, VRS, and scale efficiency measures. The visual representation shows whether the distributions are bell-shaped and provide indication about their respective
skewness. These results suggest that the use of parametric tests, i.e., normality assumptions (normality and equal variances) for the pooled $\hat{\delta}$ of CRS, VRS, and scale efficiency measures are not valid.

Figure 2: Percentage change of the DEA efficiency measures under the CRS assumption with and without a clustering approach.

Figure 3: Percentage change of the DEA efficiency measures under the VRS assumption with and without a clustering approach.
Given the distributions of $\hat{\delta}_{it}$ are non-normal, to evaluate whether the efficiency measures of the cluster-adjusted DEA model are better off in comparison to the DEA model without a clustering approach, four non-parametric tests are conducted: Kolmogorov Smirnov (KS) Statistics, Kruskal-Wallis (KW), Wilcoxon Rank Sum (WRS), and Ansari-Bradley. Without loss of generosity of the DEA efficiency measures estimated under the CRS, VRS, and scale assumptions, the null and alternative hypotheses associated with KS Statistics, KW, WRS, and Ansari-Bradley tests can be written as:

1. Kolmogorov-Smirnov (KS) Statistics Test
   
   $H_0$: There exists no difference in distribution, $D$, of the DEA efficiency measures estimated with a clustering approach, $\hat{\mu}_{ct}$, and without a clustering approach, $\hat{\mu}_{it}$. That is: $D_{\hat{\mu}_{ct}} = D_{\hat{\mu}_{it}}$.
   
   $H_a$: There exists a difference in distribution, $D$, of the DEA efficiency measures estimated with a clustering approach, $\hat{\mu}_{ct}$, and without a clustering approach, $\hat{\mu}_{it}$. That is: $D_{\hat{\mu}_{ct}} \neq D_{\hat{\mu}_{it}}$.

2. Kruskal-Wallis (KW) Test
   
   $H_0$: The difference in population median, $M$, of the DEA efficiency measures estimated with a clustering approach, $\hat{\mu}_{ct}$, and without a clustering approach, $\hat{\mu}_{it}$ is equal to zero. That is: $M_{\hat{\mu}_{ct}} - M_{\hat{\mu}_{it}} = 0$.
   
   $H_a$: Population median, $M$, of the DEA efficiency measures obtained with a clustering approach, $\hat{\mu}_{ct}$, is greater than the population median, $M$, of the DEA efficiency measures obtained without a clustering approach, $\hat{\mu}_{it}$. That is: $M_{\hat{\mu}_{ct}} - M_{\hat{\mu}_{it}} > 0$.

3. Wilcoxon Rank Sum (WRS) Test
   
   $H_0$: The difference in population median, $M$, of the DEA efficiency measures estimated with a clustering approach, $\hat{\mu}_{ct}$, and without a clustering approach, $\hat{\mu}_{it}$ is equal to zero. That is: $M_{\hat{\mu}_{ct}} - M_{\hat{\mu}_{it}} = 0$.
   
   $H_a$: Population median, $M$, of the DEA efficiency measures obtained with a clustering approach, $\hat{\mu}_{ct}$, is greater than the population median, $M$, of the DEA efficiency measures obtained without a clustering approach, $\hat{\mu}_{it}$. That is: $M_{\hat{\mu}_{ct}} - M_{\hat{\mu}_{it}} > 0$. 

![Figure 4: Percentage change of the DEA efficiency measures under the scale assumption with and without a clustering approach.](image)
4. Ansari-Bradley Test

\( H_0: \) The difference in population dispersion, \( \sigma \), of the DEA efficiency measures obtained with a clustering approach, \( \hat{\mu}_{it}^c \), and without a clustering approach, \( \hat{\mu}_{it} \), is equal to zero. That is: \( \sigma_{\hat{\mu}_{it}^c} - \sigma_{\hat{\mu}_{it}} = 0 \).

\( H_a: \) Population dispersion parameters, \( \sigma \), of the DEA efficiency measures obtained with a clustering approach, \( \hat{\mu}_{it}^c \), is greater than the population dispersion parameters, \( \sigma \), of the DEA efficiency measures obtained without a clustering approach, \( \hat{\mu}_{it} \). That is: \( \sigma_{\hat{\mu}_{it}^c} - \sigma_{\hat{\mu}_{it}} > 0 \).

Table 7 presents the nonparametric results of the differences in the pooled DEA efficiency measures under CRS, VRS, and scale assumptions estimated with and without a clustering approach. From Table 7, it can be observed that the null hypothesis is statistically rejected under KS Statistics, KW, and WRS, and accepted under Ansari-Bradley test at a 1% significance level. Hence, we conclude the following: 1) With the KS test, it exists a statistical difference in distribution of the DEA efficiency measures estimated with and without a clustering approach; 2) With KW and WRS tests, the population median of the DEA efficiency measures obtained with a clustering approach is statistically greater than the population median of the DEA efficiency measures estimated without a clustering approach; and 3) With Ansari-Bradley test, the population dispersion parameter of the DEA efficiency measures obtained with clustering approach is greater than the population dispersion parameter of the DEA efficiency measures obtained without a clustering approach. Like Table 7, Table 8 presents the nonparametric statistical results by year.

**Table 7: Nonparametric statistical tests (Pooled).**

| criteria | KS | Kruskal Wallis | Wilcoxon Rank Sum | Ansari-Bradley |
|----------|----|---------------|------------------|---------------|
| Efficiency measures under CRS assumption | | | | |
| p-value | 0.0001** | 0.0001** | 0.0001** | 0.0824 |
| Efficiency measures under VRS assumption | | | | |
| p-value | 0.0001** | 0.0001** | 0.0001** | 0.0342* |
| Efficiency measures under scale assumption | | | | |
| p-value | 0.0001** | 0.0001** | 0.0001** | 0.0566 |

** denotes the significance at a 1 percent level and * denotes the significance at a 5 percent level.
Table 8: Nonparametric statistical tests by year.

| Year | KS   | KW   | WRS   | Ansari-Bradley | Year | KS   | KW   | WRS   | Ansari-Bradley |
|------|------|------|-------|---------------|------|------|------|-------|---------------|
| 2000 | 0.0001** | 0.0001** | 0.0001** | 0.4389 | 2009 | 0.6604 | 0.2567 | 0.1289 | 0.4415 |
| 2001 | 0.001**  | 0.0057** | 0.0029** | 0.4340 | 2010 | 0.0025** | 0.0011** | 0.0005** | 0.1747 |
| 2002 | 0.0001** | 0.0001** | 0.0011** | 0.0229 | 2011 | 0.2656 | 0.1446 | 0.0727 | 0.1887 |
| 2003 | 0.0763 | 0.3474 | 0.1744 | 0.1197 | 2012 | 0.1741 | 0.0599 | 0.0301* | 0.2354 |
| 2004 | 0.0162* | 0.0309* | 0.0156* | 0.0414* | 2013 | 0.0011** | 0.0002** | 0.0011** | 0.3545 |
| 2005 | 0.0001** | 0.0001** | 0.0001** | 0.0315* | 2014 | 0.1140 | 0.0271* | 0.0136* | 0.2321 |
| 2006 | 0.1813 | 0.0672 | 0.0338* | 0.3084 | 2015 | 0.0137* | 0.0121* | 0.0006** | 0.4272 |
| 2007 | 0.0001** | 0.0001** | 0.0001** | 0.4062 | 2016 | 0.0681 | 0.1881 | 0.0945 | 0.0682 |
| 2008 | 0.6604 | 0.7727 | 0.3874 | 0.0265* | 2017 | 0.0090** | 0.0704 | 0.0354* | 0.3234 |

Efficiency measures under VRS assumption

| Year | KS   | KW   | WRS   | Ansari-Bradley |
|------|------|------|-------|---------------|
| 2000 | 0.0001** | 0.0001** | 0.0001** | 0.0350* |
| 2001 | 0.4874 | 0.2249 | 0.1130 | 0.3248 |
| 2002 | 0.0040* | 0.0015** | 0.0007** | 0.0154* |
| 2003 | 0.0049** | 0.0004** | 0.0002** | 0.2455 |
| 2004 | 0.5095 | 0.5103 | 0.2560 | 0.2345 |
| 2005 | 0.0470* | 0.0147** | 0.0074** | 0.2956 |
| 2006 | 0.5095 | 0.6100 | 0.3059 | 0.0964 |
| 2007 | 0.0001** | 0.0001** | 0.0001** | 0.0067** |
| 2008 | 0.4095 | 0.8447 | 0.4234 | 0.0264* |

Efficiency measures under scale assumption

| Year | KS   | KW   | WRS   | Ansari-Bradley |
|------|------|------|-------|---------------|
| 2000 | 0.1528 | 0.1082 | 0.0544 | 0.3504 |
| 2001 | 0.0002** | 0.0001** | 0.0001** | 0.0540 |
| 2002 | 0.0001** | 0.0001** | 0.0001** | 0.2648 |
| 2003 | 0.2656 | 0.4592 | 0.2304 | 0.3924 |
| 2004 | 0.1196 | 0.0318* | 0.0160* | 0.1858 |
| 2005 | 0.0006** | 0.0014** | 0.0007** | 0.0028** |
| 2006 | 0.2656 | 0.1145 | 0.0575 | 0.1333 |
| 2007 | 0.0470* | 0.0206* | 0.0104* | 0.3501 |
| 2008 | 0.5095 | 0.4638 | 0.2327 | 0.4311 |

** denotes the significance at a 1 percent level and * denotes the significance at a 5 percent level.

6 Challenges and Conclusions

This paper, addressing the issues associated with extreme data points and heterogeneity found in the linear programming data envelopment analysis (DEA) model, presents an alternative cluster-adjusted DEA model. However, unlike existing literature that defines the clusters based on inputs-outputs, we define the clusters based on the DEA efficiency measures. The number of clusters based on the DEA efficiency measures is statistically determined using Gap statistic and Elbow methods. We use the December quarterly panel data consisting of 122 U.S agricultural banks across 37 states from 2000 to 2017 to estimate the cluster-adjusted DEA model.

The proposed cluster-adjusted DEA model involves 4 stages (or steps). First, we estimate the efficiency measures using linear programming DEA model. Second, based on the estimated efficiency measures, the optimal number of cluster groups is determined using the Gap statistic and Elbow methods. These results are further validated by the 30 indices of the NbClust package. Accordingly, the majority rule of the clustering indices concluded that the optimal number of clusters is three groups. Furthermore, these results were supported by the distribution of DEA efficiency measures under the CRS assumption (Figure 1).

Third, using the statistically identified clusters of banks, we estimate the cluster-adjusted DEA model
while accounting for the yearly variability. Finally, in the evaluation of differences in the efficiency measures estimated with the DEA and cluster-adjusted DEA models, the nonparametric tests of Kolmogorov-Smirnov Statistics, Kruskal-Wallis, Wilcoxon Rank Sum, and Ansari-Bradley are conducted. These tests were conducted to compare the distributions, medians, and dispersions of the DEA and cluster-adjusted DEA efficiency estimators. Our results provide evidence that the deterministic DEA model does not guarantee accurate efficiency measures in the presence of non-homogeneous banks or DMUs.

However, there are limitations that future researchers could study to improve the discriminatory power of the cluster-adjusted DEA results. For example, future research could incorporate banks merger and acquisition in order to achieve optimal economies of scale. Compared to the current framework, the results of the efficiency measures could vary regarding the type of mergers and the number of yearly mergers. It could be great to further incorporate the financial crisis as a dummy and study its implication of the DEA cluster efficiency measures. Research could also focus on classifying the total assets using the FCA’s classification requirement and comparing the statistical properties with the clustering approach of DEA efficiency measures.

Financial Support: This research was supported by the Center for Agricultural Policy and Trade Studies, North Dakota State University [N/A].

Conflict of Interests Statement: The authors have no conflicts of interest to disclose.

References

Alirezaee, M., & Sani, M. (2011). New analytical hierarchical process/data envelopment analysis methodology for ranking decision-making units. *International Transactions in Operational Research*, 18(5), 533-544. https://doi.org/10.1111/j.1475-3995.2011.00812.x

Aly, H., Grabowski, R., Pasurka, C., & Rangan, R. (1990). Technical scale and allocative efficiencies in US banking: An empirical investigation. *Review of Economics and Statistics*, 72 (2), 211-218. https://doi.org/10.2307/2109710

Amirteimoori, A., & Kordrostami, S. (2014). Data envelopment analysis with discrete-valued inputs and outputs. *Expert Systems*, 31(4), 335-342. https://doi.org/10.1111/exsy.12045

Atwood, J., & Shaik, S. (2018). Quantile DEA: Estimating qDEA-alpha Efficiency Estimates with Conventional Linear Programming. In: *Greene W., Khalaf L., Makdissi P., Sickles R., Veall M., Voia MC. (eds) Productivity and Inequality*. Springer Proceedings in Business and Economics. Springer, Cham. https://doi.org/10.1007/978-3-319-68678-3_14

Banker, R., Charnes, A., & Cooper, W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30(9), 1078-1092. https://doi.org/10.1287/mnsc.30.9.1078

Banker, R., & Chang, H. (2006). The super-efficiency procedure for outlier identification, not for ranking efficient units. *European Journal of Operational Research*, 175 (2), 1311-1320. https://doi.org/10.1016/j.ejor.2005.06.028

Casu, B., & Molyneux, P. (2003). A comparative study of efficiency in European banking. *Applied Economics*, 35, 1865-1876. https://doi.org/10.1080/0003684032000158109

Charnes, A., Cooper, W., Lewin, A., & Seiford, L. (1994). Data envelopment analysis: theory, methodology and application. *Boston: Kluwer Academic Publishers.*

Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of DMUs. *European Journal of Operational Research*, 2, 429-444. https://personal.utdallas.edu/~ryoung/phdseminar/CCR1978.pdf

Chen, T. (2002). A comparison of chance-constrained DEA and stochastic frontier analysis: Bank efficiency in Taiwan. *The Journal of the Operational Research Society*, 53(5), 492-500. https://doi.org/10.1057/palgrave/jors/2601318

Crowley, J. (2007). Interest Rate Spreads in English-Speaking Africa. *International Monetary Fund*. https://doi.org/10.5089/9781451866650.001

Dai, X., & Kuosmanen, T. (2014). Best-practice benchmarking using clustering methods: Application to energy regulation. *Omega*, 42(1), 179-188. https://doi.org/10.1016/j.omega.2013.05.007

Dipayan, R. (2014). Analysis of Technical Efficiency of Indian Banking Sector: An Application of Data Envelopment Analysis. *International Journal of Finance and Banking Studies*, 3(1), 150-160. https://ideas.repec.org/a/rbs/ijfbss/v3y2014i1p150-160.html

Drake, L., & Hall, M. (2003). Efficiency in Japanese banking: An empirical analysis. *Journal of Banking and Finance*, 27(5), 891-917. https://doi.org/10.1016/S0378-4266(02)00240-6

Emrouznejad, A., Parker, B., & 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-Economics Planning Science*, 42(3), 151-157. https://doi.org/10.1016/j.
Farrell, J. (1957). The Measurement of productive efficiency. *Journal of Royal Statistical Society*, Series A (120): 253-290. https://doi.org/10.2307/2934100

Farrell, J., & Fieldhouse, M. (1962). Estimating efficient production under increasing returns to scale. *Journal of the Royal Statistical Society, Series A, General*, 125(2), 252-267. https://doi.org/10.2307/2982329

Flannery, M. (1980). How do changes in market interest rates affect bank profits? *Business Review*, Federal Reserve Bank of Philadelphia. https://www.philadelphiafed.org/-/media/research-and-data/publications/business-review/1980/br80somf.pdf

Gardner, M., Mills, D., & Cooperman, E. (2005). Managing financial institutions. 5th Edition. *McGraw-Hill Higher Education*

Hauner, D. (2005). Explaining efficiency differences among large German and Austrian banks. *Applied Economics*, 37, 969-980. https://doi.org/10.1080/00036840500081820

Jahangoshai, R., Jozmaleki, M., & Valipour, M. (2018). Integrating dynamic fuzzy C-means, data envelopment analysis and artificial neural network to online prediction performance of companies in stock exchange. *Physica A: Statistical Mechanics and Its Applications*, 489, 78-93. https://doi.org/10.1016/j.physa.2017.07.017

Jessen, A. (2009). A multi-criteria block model for performance assessment. *Omega*, 37, 204-214.

Koopmans, C. (1951). Activity Analysis of Production and Allocation.

Li, S., Jahanshahloo, G., & Khodabakhshi, M. (2007). A super-efficiency model for ranking efficient units in data envelopment analysis. *Applied Mathematics and Computation*, 184, 638-648. https://doi.org/10.1016/j.amc.2006.06.063

Maletic, R., Maletic, P., Krecz, M., & Popovic, B. (2013). Comparative analysis of ranking municipalities in Vojvodina using DEA and I-distance methods. *International Journal of Agricultural and Statistical Sciences*, 9(2), 471-480.

Meimand, R., Cavana, Y., & Laking, R. (2002). Using DEA and survival analysis for measuring performance of branches in New Zealand’s accident compensation corporation. *Journal of Operational Research Society*, 53(3), 303-313. https://doi.org/10.1057/palgrave.jors.2601291

Miller, M. & Noulas, G. (1996). The technical efficiency of large bank production. *Journal of Banking and Finance*, 20(3), 495-509. https://doi.org/10.1016/0378-4266(95)00017-8

Paradi, I., Zhu, H., & Edelstein, B. (2012). Identifying managerial groups in a large Canadian bank branch network with a DEA approach. *European Journal of Operational Research*, 219, 178-187. https://doi.org/10.1016/j.ejor.2011.12.022

Po, Y., Guh, Y., & Yang, M. (2009). A new clustering approach using data envelopment analysis. *European Journal of Operational Research*, 199, 276-284. https://doi.org/10.1016/j.ejor.2008.10.022

Saatli, S., Hatami-Marbini, A., Tavana, M., & Aghajani, P. (2013). A fuzzy data envelopment analysis for clustering operating units with imprecise data. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 21(1), 29-54. https://doi.org/10.1142/S0218478X13500373

Sakouvogui, K. (2020). A Comparative clustering approach of Stochastic Frontier Analysis and Data Envelopment Analysis estimators: Evidence from Banking System. *Journal of Economics studies*. Forthcoming

Samioilenko, S., & Osei-Bryson, K. (2008). Increasing the discriminatory power of DEA in the presence of the sample heterogeneity with cluster analysis and decision trees. *Expert Systems with Applications*, 34(2), 1568-1581. https://doi.org/10.1016/j.eswa.2007.01.039

Sengupta, J. (1992). A fuzzy systems approach in data envelopment analysis. *Computers and Mathematics with Applications*, 24(8-9), 259-266. https://doi.org/10.1016/0898-1221(92)90203-T

Shark, S. (2013). Crop Insurance Adjusted Panel Data Envelopment Analysis Efficiency Measures. *American Journal of Agricultural Economics*, 95(5), 1136-1154. https://doi.org/10.1093/ajae/aat062

Shark, S., Mishra, A., & Atwood, J. (2012). Aggregation issues in the estimation of linear programming productivity measures. *Journal of Applied Economics*, 15(1), 169-187. https://doi.org/10.1080/1540848120260008-7

Sugan, C. (1998). Techniques for Clustering and Classification with Applications to Medical Problems. Ph.D. Dissertation, *Department of Statistics, Stanford University*, Stanford, California, United States.

Shokouhi, A., Hatami-Marbini, M., Tavana, M., & Saati, S. (2010). A robust optimization approach for imprecise data envelopment analysis. *Computers and Industrial Engineering*, 59(3), 387-397. https://doi.org/10.1016/j.cie.2010.05.011

Tao, L., Liu, X., & Chen, Y. (2013). Online banking performance evaluation using data envelopment analysis and axiomatic fuzzy set clustering. *Quality and Quantity*, 47(2), 1259-1273. https://doi.org/10.1007/s11135-012-9767-3

Thanassoulis, E. (1996). A data envelopment analysis approach to clustering operating units for resource allocation purposes. *Omega*, 24(4), 463-476. https://doi.org/10.1016/0305-0483(96)00013-8

Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the Number of Clusters in a Data Set Via the Gap Statistic. *Journal of the Royal Statistical Society B*, 63(2), 411-423. https://doi.org/10.1111/1467-9868.00293

Tone, K. (2017). Advances in DEA Theory and Applications: with extensions to forecasting models. *John Wiley and Sons Inc*, New York, United States

Vittas, D. (1991). Measuring commercial bank efficiency: Use and misuse of bank operating ratios. *Policy Research working paper Series*, 806, World Bank
Wu, D., Yang, Z., & Liang, L. (2006). Using DEA-neural network approach to evaluate branch efficiency of a large Canadian bank. *Expert Systems with Applications*, 31(1), 108-115. https://doi.org/10.1016/j.eswa.2005.09.034

Yang, T., & Kuo, C. (2003). A hierarchical AHP/DEA methodology for the facilities layout design problem. *European Journal of Operational Research*, 147(1), 128-136. https://doi.org/10.1016/S0377-2217(02)00251-5