Scientific
Research
Publishing

An Examination of Very Low Efficiency Scores in Data Envelopment Analysis in the Pension Funds Industry

Maryam Badrizadeh*, Joseph C. Paradi

Centre for Management of Technology and Entrepreneurship, University of Toronto, Ontario, Canada
Email: *maryam.badrizadeh@mail.utoronto.ca

Abstract

Data Envelopment Analysis (DEA) is a powerful analytical tool that is considered as one of the most useful techniques to measure the efficiency of Decision Making Units (DMUs) in certain industry segments. However, there is a scarcity of reported use to assess pension funds' performance due to the complexities of such funds. The few papers that can be found in literature do not consider the main characteristics of pension funds such as uncontrollable variables for managers, regulations, and funds' status (fully funded/underfunded pension plans). Regulations affect such investment vehicles in many ways from investment strategy, tax status, reporting requirements and others. Also, as the by-product of our past research in this field the authors ran into some unexpected outcomes where some funds had achieved an extremely low efficiency score. This is very highly unusual and invited additional research. There are very few papers in the literature on extremely low efficiency scores, and there is a paucity of cogent explanations on why this is the case. Therefore, while evaluating the pension funds' performance through DEA we worked on this problem in some detail to uncover the reason(s) for such low minimum efficiency scores for pension funds. We found that the presence of very low efficiency scores phenomena is not uncommon in pension funds industry but is in other industry studies.

1. Introduction[1]

DEA points out the efficient DMUs which can be considered as a target for the inefficient ones (Charnes et al., 1978). Therefore, DEA is widely used in the literature for banking corporations, branch networks, risk, fraud and many other financial institutions. However, there are only a few papers that used DEA to evaluate pension funds' performance. In 2005, Barrientos and Boussofiane studied the efficiency of pension fund managers in Chile by using CCR and BCC models in DEA for the period of 1982-1999. The results indicated that the Chilean companies exhibited significant inefficiency. There were changes over time but no continuous trends toward an efficiency improvement (Barrientos & Boussofiane, 2005). In 2006, Barros and Garcia evaluated Portuguese pension funds' performance from 1994-2003 by using different DEA models such as CCR, BCC, Cross-Efficiency and Super-Efficiency and compared the results. The results indicated that the majority of Portuguese firms displayed relatively high managerial skills, being VRS-efficient for the most part. However, there were some inefficient firms which could have improved much more. Also, their results showed that private pension funds managers were more efficient than those of public institutions. The researchers also tested institutions involved in mergers and acquisitions during the period and found that they were more efficient than those that were not involved in these processes. The results supported that small pension funds management companies which did not merge, had less efficient performance and their size disadvantage acted as a negative influence on their operations. Unfortunately, they studied only 12 pension funds and the authors cautioned the readers about this problem (Barros & Garcia, 2006). DEA requires a sufficient number of observations to allow good separation and discrimination amongst DMUs. A small sample size can reduce the accuracy of results. The rule of thumb is that the number of DMUs should be at least three times the total number of inputs plus outputs which are used in the model (Cooper et al., 2011). In 2010, Garcia analyzed changes in the productivity of the same 12 Portuguese pension funds management institutions for a different timeline of 1994-2007. He used DEA and the Malmquist index. His results indicated that increasing the governance and transparency of the fund's management companies would increase their efficiency (Garcia, 2010). In 2011, Sathye estimated the production efficiency of pension funds in Australia for the years 2005 to 2009 using CCR and BCC models. The results indicated that Australia's pension funds' efficiencies were too low in some cases. They carried out regression analyses on the variables and found that certain fund characteristics, such as size and the pro-

---

[1]Note that some of the following parts had appeared in book chapters by the authors of this paper: Badrizadeh, M., "Pension Funds Insights with DEA", Book Chapter in "Data Envelopment Analysis in the Financial Services Industry. A Guide for Practitioners and Analysts Working in Operations Research Using DEA International Series in Operations Research & Management Science". Paradi, J. C., Sherman H. D., Tam F. K. Springer, 2018; and Badrizadeh, M., Paradi, J. C., "Mixed Datasets with Partially Deficient Variable Sets Embodied in Mixed Variable DEA (MV-DEA)", Book Chapter in "Data Science and Productivity Analytics", Charles, V., Aparicio, J., and Zhu, J., Springer, 2020.

portion of funds invested in low-risk opportunities had a positive association with performance. They also discovered that diversification and the financial crises, as one would expect, yielded a negative effect on efficiency (Sathye, 2011). In 2015, Galagedera and Watson assessed pension funds in Australia by using DEA for year 2012. In the study the funds were classified under four categories: industry, public sector, corporate and retail. The results showed that retail funds were the best performers (Galagedera & Watson, 2015). However, each of these categories has its own specific characteristics and it might not be proper to consider all of them in one model without providing certain specific treatments. In 2017, Badrizadeh explained the importance of considering different characteristics of investment industries by comparing pension funds and mutual funds (Badrizadeh, 2017). In 2015, Zamuee evaluated Namibian pension funds using a CCR model for years 2010 to 2014. The results showed that the majority of these pension funds were found to have low efficiency scores and urgent management intervention was required to address the issues raised (Zamuee, 2015).

The previous studies on pension funds using DEA predominantly focused on comparing different DEA models instead of having a clear methodology and framework. Moreover, as mentioned above, DEA requires sufficient sample size to allow good separation and discrimination amongst DMUs. Most of these studies had very few DMUs considering the number of inputs and outputs which not only distorts their results, but the findings are to be treated with caution. Also, a major flaw of these studies is that none of them consider the effect of government regulations on these pension plans' performance that significantly restrict and impact the managers' control and differentiate them from other investment vehicles that do not have such restrictions. For instance, contribution amounts and benefit payments are the main variables here. However, there are government regulations for these two variables, and they are not completely under the managers' control. Furthermore, one of the important issues in the pension funds industry is their financial health with respect to their ability to meet their legal obligations to see whether pension plans are fully funded, or they have deficits and are underfunded. Therefore, fully funded plans are referenced only to their own category while the underfunded plans are referenced to their level as well as the higher level of hierarchy (fully funded plans). None of these studies considered these important characteristics of pension funds which distinguish them from other investment vehicles. As Badrizadeh and Paradi cautioned, these factors should be treated appropriately (Charles et al., 2020).

This research investigates the performance of different private pension plans by considering the effects of regulations on asset allocation as well as managers' authority on the contribution amounts and benefit payments as the two main variables in this industry, plus there are other factors. The issue in focus here is the very low efficiency/productivity results in the pension funds industry compared to other industries such as:

- Agriculture: 0.2 to 0.4 (Blancard & Martin, 2013) and (Vlontzos et al., 2014)

- Airlines: 0.1 to 0.3 (Barros & Peypoch, 2009) and (Wanke & Barros, 2016)
- Banking: 0.4 to 0.6 (Asmild et al., 2004) and (Chiu et al., 2008)
- Mutual Funds: 0.1 to 0.3 (Basso & Funari, 2005) and (Premachandra et al., 2012)
- Pension Funds: 0.00 to 0.008 (Sathye, 2011) and (Zamuee, 2015)

The rest of the article is organized as follows: the methodology is explained in Section 2 and the results are discussed in Section 3. The conclusions are presented in Section 4.

2. Methodology

The problem of low efficiency for DMUs arose from work we did on efficiency comparisons between federally regulated pension funds and Canadian Mutual Funds (Badrizadeh, 2017). In this section, first the main characteristics of pension funds are investigated, calculated, and considered in the model. Then, the reason for low minimum efficiency scores for pension funds is investigated.

2.1. Evaluating Pension Funds' Performance

Canadian pension plans are mostly categorized into two groups: Defined Benefit Plans (DB) and Defined Contribution Plans (DC). DB plans offer an employee the security of knowing what to expect at retirement based on their salary during their working years. In a DB plan, the investment risk of guaranteed retirement income is taken on by the employer. DC plans specify the amount of employer and employee contributions. The amount available to provide a pension income in a DC plan is affected by how successfully contributions are invested. Therefore, the investment risk in a DC plan rests with the employee. The plan members who are not covered by DB or DC plans, are covered by other types of plans such as combination pension plans (Combo). Combo plans incorporate both defined benefit and defined contribution plans' concepts. This type of plan offers additional flexibility for plan members and employers by incorporating some positive elements from both plans. In a combination plan, a pension must be promised and from time to time, employers are able to use pension surplus to fund their defined benefit plan's current service costs. In this research, these three types of plans are studied.

One of the important issues in managing such funds is government regulations intended to protect the retirement income and maximize returns. Pension laws and regulations shape the unique legal investment environment in which pension funds operate. Managers consider the impact of regulations as a matter of course in their work. In general, regulations can be categorized into two types. One type of regulation deals with administration of the various types of pension plans. These rules are numerous and change from one situation to another based on actuarial age considerations, mortality, conditions on fund transfer to spouse/common-law partner after death, etc. The second type of regulation deals with the allocation of assets. For instance, in Canada, according to the Financial Ser-

vices Commission of Ontario in compliance with Federal Investment Regulations rules, a maximum of 5% of the plan's assets may be invested directly or indirectly in any Canadian resource property. Similarly, a maximum of 10% of the assets may be invested in or loaned to any one person/associated person[2] (FSCO, 2004). Hence, the only part of the regulations that managers can somehow control is how to allocate their assets to different investment vehicles more effectively while observing these restrictions. As a result, if the effects of regulations on asset allocation can be quantified, then the managers' performance for allocating their assets can be assessed. In order to achieve this goal, the standard deviation of returns is calculated based on their asset allocation and added to the variables. As maximization of the benefits is an important objective, a Variable Returns to Scale (VRS) output-oriented model was chosen as the base model. The VRS model was proposed by Banker, Charnes and Cooper in 1984 (Banker et al., 1984). The VRS frontier does not pass through the origin as the Charnes, Cooper and Rhodes model (CCR) does. But here, a few of the variables are not completely under fund-managers' control hence, these should be treated accordingly. Some variables such as contribution amounts and benefit payments which are also not under the managers' control, are considered as non-discretionary variables in the model. Consequently, such variables are removed from the objective function of the linear program but are included in the constraints to assure their influence and their values remain constant while the discretionary variables are optimized (Banker & Morey, 1986).

Let's consider DMUs ($i = 1, 2, \cdots, n$), inputs ($j = 1, \cdots, m$) and outputs ($r = 1, \cdots, s$). The "D" and "ND" refer to the "Discretionary" and "Non-Discretionary" input and output. The mathematical model for multiplier form of the Non-Discretionary VRS (Non-Dis-VRS) is presented in Equation (1):

$$\min \ z = \sum_{j \in D} v_j x_{jo} + \sum_{j \in ND} v_j x_{jo} - \sum_{r \in ND} u_r y_{ro} - v_o \tag{1}$$

Subject to: $\sum_{j \in D} v_j x_{ji} + \sum_{j \in ND} v_j x_{ji} - \sum_{r \in ND} u_r y_{ri} - \sum_{r \in D} u_r y_{ri} - v_o \geq 0$

$$\sum_{r \in D} u_r y_{ro} = 1$$

$$v_j \geq \varepsilon, \ \ j \in D$$

$$v_j \geq 0, \ \ j \in ND$$

$$u_r \geq \varepsilon, \ \ r \in D$$

$$u_r \geq 0, \ \ r \in ND$$

The envelopment form of the Non-Dis-VRS model is shown in Equation (2):

$$\max \ \varphi + \varepsilon \left(\sum_{j \in D} s_j + \sum_{r \in D} t_r \right) \tag{2}$$

Subject to: $\sum_{i=1}^{n} \lambda_i y_{ri} = t_r + \varphi y_{rio}, \ \ r \in D$

$$\sum_{i=1}^{n} \lambda_i y_{ri} = t_r + y_{rio}, \quad r \in ND$$

$$\sum_{i=1}^{n} \lambda_i x_{ji} = -s_j + x_{jio}, \quad j = 1, \cdots, m$$

$$\sum_{i=1}^{n} \lambda_i = 1$$

Moreover, in order to have a clear insight into how the pension funds industry is managed, one has to consider the funds' status to see whether the plans meet their financial and actuarial obligations, hence fully funded or they have deficits, and these are underfunded. Therefore, the categorical variables are also considered in the Non-Dis-VRS model and signal whether an under/fully funded condition exists. To achieve this goal, fully funded plans are referenced only to their own category while the underfunded plans are referenced to other, similar plans as well as to the fully funded ones. It should be noted that all pension plans for this research are active plans. Therefore, if an active plan is flagged as underfunded it means that the plan will not meet its obligations based on the government tiers' deadlines (generally in the next 5 or 10 years). During these years, if the underfunded active plan's managers compensate for the financial deficits, the plan should move to the fully funded category. If not, after the deadline (5 or 10 years) the plan may be terminated. Only a few underfunded active plans are found efficient.

2.2. Investigating Low Minimum Efficiency Scores

Only DB and Combo plans' managers must file benefit payments in their annual reports to the regulator(s) while this is optional for DC plans. The dataset in this article covers 2010. According to the Office of the Superintendent of Financial Institutions Canada (OSFI), in 2011, only about 10% of the plans under federal regulations were fully funded while the rest were underfunded[3]. In order to have a comprehensive study of this industry, both fully funded and underfunded plans were considered.

From other studies using DEA, only two, Sathye (2011) and Zamuee (2015), mentioned the minimum efficiency scores problem we had found. They used various DEA models and different years and found some efficiency scores in the range of 0.00 to 0.08. As evident in these two articles as well as the results of this study, low minimums do occur in the pension funds industry.

Therefore, a hypothesis can be made as follows:

Hypothesis A: Since both fully funded and underfunded active pension plans should be considered in the DEA model and most active pension plans are underfunded, very low minimum efficiency scores could be found in such studies.

In order to investigate this hypothesis, three different approaches were carried out. In the first approach, the efficient DMUs are removed from the analysis. The efficient DMUs were divided as fully funded and underfunded plans. However, the former can be only referenced to like plans but the latter can be refe-

---

[3]http://www.osfi-bsif.gc.ca/Eng/Pages/default.aspx; and
http://www.thewhig.com/2012/10/05/most-major-pension-plans-underfunded-supervising-body.

renced to either or both. The efficient DMUs are "peeled off" from the frontier until there are no more fully funded plans remain and the average and minimum efficiency scores are examined. In the second approach, the inefficient DMUs are counted to see how many of them are from underfunded plans. In the third approach, the models are run for fully funded and underfunded plans separately to check the minimum efficiency scores.

A detailed explanation of the data and the results of the Non-Dis-VRS model is presented in Section 3.

3. Results and Discussion

3.1. Data Preparation

Data from all Canadian DB plans, DC plans and Combo plans which are supervised federally and not terminated during year 2010 and have at least 100 or more plan members were studied. Although OSFI is the best source for pension funds data, extensive work was performed to validate the data. Moreover, standard deviation of returns which is an important variable was not provided by OSFI. Therefore, for each pension plan this was calculated from the available data for asset allocation. Also, outliers which would have unwanted effects on the results of a DEA model were removed. Therefore, the variables were selected carefully and validated through various techniques such as statistical tests, sensitivity analysis and outlier removal which are explained below.

3.1.1. Number of DMUs

DEA requires a sufficient number of DMUs to allow good separation and discrimination amongst them. A small sample size can reduce the accuracy of the results. As explained earlier, the rule of thumb is that the number of DMUs should be at least three times the total number of inputs plus outputs which are used in the model. Another similar rule is: $n \geq \max\{m \times s, 3 \times (m + s)\}$, where n is the number of DMUs, m is the number of inputs and s is the number of outputs (Cooper et al., 2011). As presented later in this section, there are 4 inputs and 2 outputs for DB plans and Combo plans and 3 inputs and 1 output for DC plans. There are 90 DB plans, 37 DC plans and 46 Combo plans. Therefore, the number of pension plans is sufficient for this research.

3.1.2. Variable Selection
1) Efficiency Contribution Analysis

In this method, an input or output variable is included or excluded from a DEA model in order to determine the influence of each variable on the DEA scores (Smith, 1997). Therefore, first, each variable is removed, and the model is re-run with the rest of the variables. Then, the difference in average scores between the original model and the re-run model are compared to determine whether the variable has a significant impact on the DEA scores.

2) Correlation Analysis

The correlation coefficient used here which represents the relationship between

two variables is known as the Pearson correlation coefficient and can be calculated as shown in Equation (3):

$$\rho_{x,y} = \frac{\text{cov}(x,y)}{\sigma_x \sigma_y} = \frac{E\left[(x-\mu_x)(y-\mu_y)\right]}{\sigma_x \sigma_y} \tag{3}$$

where x, y are the variables being compared;

 $\text{cov}(x,y)$ is the covariance of x and y;

 σ_x, σ_y is the standard deviation of x and y;

 E is the expected value;

 μ_x, μ_y is the mean of variable x and y.

The coefficient value lies between −1 and +1. A coefficient of +1 indicates the two variables are perfectly positively correlated, so as one variable increases, the other increases proportionally. Conversely, a coefficient of −1 represents a perfect negative relationship which means if one variable increases, the other decreases proportionally. A coefficient of zero demonstrates no linear relationship at all and so if one of the variables changes, the other stays the same (Field, 2009).

Correlation analysis is a common method for selecting DEA variables. The correlation coefficient can be calculated for every combination of inputs and outputs and if two inputs or outputs are highly correlated with one another, it suggests that the relationship of each variable is very close when relating it to the other. Therefore, one of the two variables may be enough and one of them can be removed. However, both management's perspective and the research's objectives should be considered in this regard.

3.1.3. Outlier Detection

1) Manual Cleaning

A simple and necessary approach for removing outliers is checking the data manually. Although this technique is very time consuming, it is essential to examine if there are any outliers in the dataset (for instance, the investment expense for the pension fund is $2). Manual data cleaning together with other methods leads to a more precise and valuable dataset. For the purpose of this section, first the data was examined manually using the histogram of each variable and the meaningless values were pinpointed. Then, the methods which are explained below were run. The results of these methods showed that indeed the manually found data were actually outliers.

2) Stripping the Efficient Frontier Approach

Stripping the frontier is an alternative method to determine whether the efficient DMUs on the frontier are, in fact, outliers (Cooper et al., 2011). In order to examine this method, all of the efficient DMUs on the frontier should be removed and the DEA model rerun. The frontier stripping is a good way to distinguish the obvious outliers on the frontier. However, it is better to consider the results of this method alongside the results of other methods.

3) Super Efficiency Test

Another approach to sensitivity analysis in DEA is removing DMUs that are

referenced by many units as they might be super-efficient outliers that skew the frontier (Simar, 2003). Therefore, the supper-efficiency model was used to determine the impact of DMUs that skew the frontier and affect the DEA scores.

3.1.4. Sensitivity Analysis

1) Wilcoxon Rank-Sum Test

The Wilcoxon Rank-Sum test is a nonparametric statistical test that examines whether the two groups belong to the same population or whether they differ significantly. It can be used when the population is not normally distributed. Since the theoretical distribution of the efficiency score in DEA is usually unknown, the Wilcoxon Rank-Sum test is a suitable method to test the DEA scores which are statistically independent (Cooper et al., 2007). For the purpose of this research, the Wilcoxon Rank-Sum test was run to check whether different pension plans can be examined in one group. The results show that only DB plans and Combo plans can be placed in one group.

According to Cooper et al., if the two independent datasets of efficiency scores are represented by $A = \{a_1, a_2, \cdots, a_m\}$ and $B = \{b_1, b_2, \cdots, b_n\}$, the combination of A and B would form C which contains $m+n$ observations in a new dataset that is arranged in descending order. Those identical values in C receive a mid-rank which means the sum of ranks divided by the number of identical values. If the sum of the ranks from A and B is S, then the normalized S would be defined as shown in Equation (4):

$$T = \frac{S - m(m+n+1)/2}{\sqrt{mn(m+n+1)/12}} \tag{4}$$

T has an approximately standard normal distribution and by using T, the null hypothesis that the two groups have the same population at a level of significance of α can be examined. The hypothesis is rejected if $T \le -T_{\alpha/2}$ or $T \ge T_{\alpha/2}$, where $T_{\alpha/2}$ corresponds to the upper $\alpha/2$ percentile of the standard normal distribution (Cooper et al., 2007).

The null hypothesis of Wilcoxon Rank-Sum Test is that the two groups have the same population. The result of the hypothesis test should be 1 or 0. If $h = 1$, this indicates rejection of the null hypothesis and if $h = 0$, this represents a failure to reject the null hypothesis at a significance level of $\alpha = 0.05$. Also, ρ-value of the test is a positive scalar from 0 to 1 with both extreme values expressing the complete separation of the distributions and 0.5 demonstrating full overlap.

In this research, the datasets were carefully examined by using all the methods mentioned above.

After data preparation, there are 173 pension plans which contain 90 DB plans, 37 DC plans, and 46 Combo plans. All pension plans must submit an annual financial statement to OSFI. For some of the variables such as benefit payments and management fees only the DB and Combo plans have to report the information, including estimates of the value of the future obligation to its employees, while this is optional for DC plans. The reporting requirements for DC plans are

straightforward since there is no future obligation for them to report, hence, the available data for DC plans is less than DB and Combo plans. Therefore, the study must be carried out in two separate parts, one part for the DB and Combo plans and one part for the DC plans. The Wilcoxon rank sum test was run to assure statistically that DB plans and Combo plans have the same population and can be considered in one group.

The inputs and outputs for DB and Combo plans are presented in Table 1. The variables are measured in Canadian dollars.

For DC plans the inputs and outputs are shown in Table 2.

Table 1. Inputs and outputs for DB Plans and combo plans.

Inputs	Outputs
• Standard Deviation of Returns	• Net Investment Income
• Investment Expenses	• Benefit Payments
• Management Fees	
• Contribution Amounts	

Table 2. Inputs and outputs for DC plans.

Inputs	Outputs
• Standard Deviation of Returns	• Net Investment Income
• Investment Expenses	
• Contribution Amounts	

3.2. Results for Evaluating Pension Funds' Performance

The output oriented Non-Dis-VRS model is used to evaluate the pension funds' performance by considering the effects of regulations on pension funds by calculating the standard deviation of returns, based on asset allocations, including controllable and uncontrollable variables. Also, the categorical variables are considered in the Non-Dis-VRS model to indicate the funds' status (fully funded and underfunded) for DB and Combo plans. Since DC plan managers are not obliged to provide defined benefit payments for their retirees, the funds' status is not an issue for DC plans. The results for the Non-Dis-VRS model for DB, Combo and DC plans are represented in Table 3 and Table 4.

The main characteristics of pension plans such as the effects of regulations on asset allocation and managers' authority as well as the funds' status are accounted for in the DEA models for the first time. As shown in Table 3 and Table 4, the average of the lowest efficiency scores' quartile for DB plans (0.17) and for DB and Combo plans (0.15) are lower than the average of the lowest efficiency scores' quartile for DC plans (0.33). Further investigation into private pension funds is carried out in Section 3.3 to examine the reasons for the very low minimum efficiency scores. This should provide a better understanding of the pension funds industry.

Table 3. Results (θs) for Non-Dis-VRS model for DB and combo plans.

Non-Dis-VRS O_CAT[a]	DB Plans	#DMUs: 90
		#Efficient: 33
		Average: 0.58
		Max: 1
		Min: 0.11
		Ave of Lowest Quartile (22 DMUs): 0.17
	DB & Combo Plans	#DMUs: 136
		#Efficient: 39 (DB:34, Combo:5)
		Average: 0.51
		Max: 1
		Min: 0.06
		Ave of Lowest Quartile (34 DMUs): 0.15

[a]Non-Dis-VRS_O_CAT: Non-Discretionary VRS, Output Oriented, Categorical DMUs (fully funded and underfunded).

Table 4. Results (θs) for Non-Dis-VRS Model for DC Plans

Non-Dis-VRS_O[a]	DC Plans	#DMUs: 37
		#Efficient: 13
		Average: 0.7
		Max: 1
		Min: 0.19
		Ave of Lowest Quartile (9 DMUs): 0.33

[a]Non-Dis-VRS_O: Non-Discretionary VRS, Output Oriented. For DC plans, the funds' status (fully funded/underfunded) is not an issue. Therefore, the categorical DMUs are not considered for the DEA model for DC plans.

3.3. Results for Investigating Low Minimum Efficiency Scores

From 90 DB plans in the dataset, 29 plans are fully funded, and the rest of the plans (61 plans) are underfunded. From 46 Combo plans, only 8 plans are fully funded and 38 plans are underfunded.

In **Figure 1** and **Figure 2** the efficiency scores' distributions are presented; sorted in descending order of their efficiency scores (θ).

As shown in **Figure 1**, approximately 30% of DMUs have θ between 0.1 to 0.3 and all these DMUs are underfunded plans.

In **Figure 2**, 54 DMUs achieved θ between 0.06 to 0.3. Most of these DMUs are from underfunded plans.

The results for the three approaches that were mentioned in the methodology section are presented below.

3.3.1. Stripping Frontiers
1) DB Plans
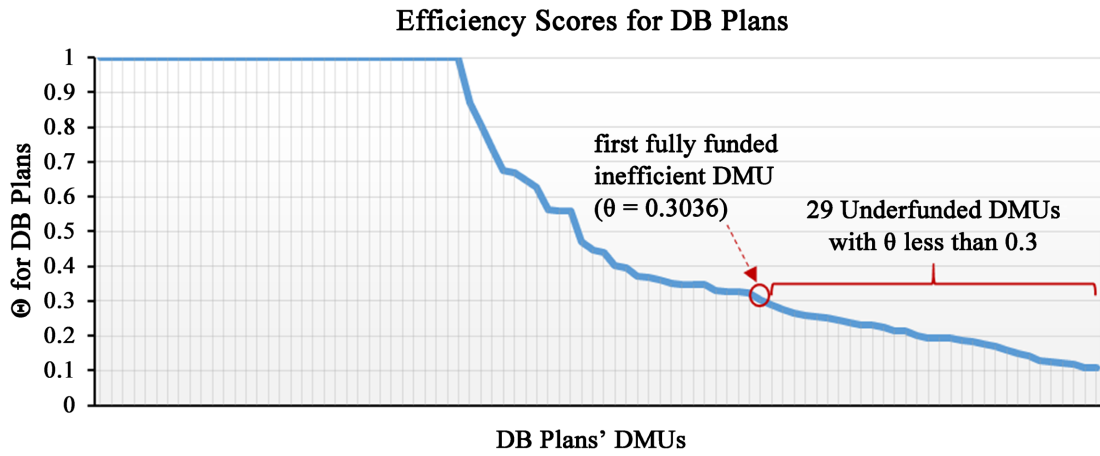There are 90 DB plans of which 29 are fully funded and 61 are underfunded:

Efficiency Scores for DB Plans

Figure 1. Efficiency scores' distribution for DB plans.

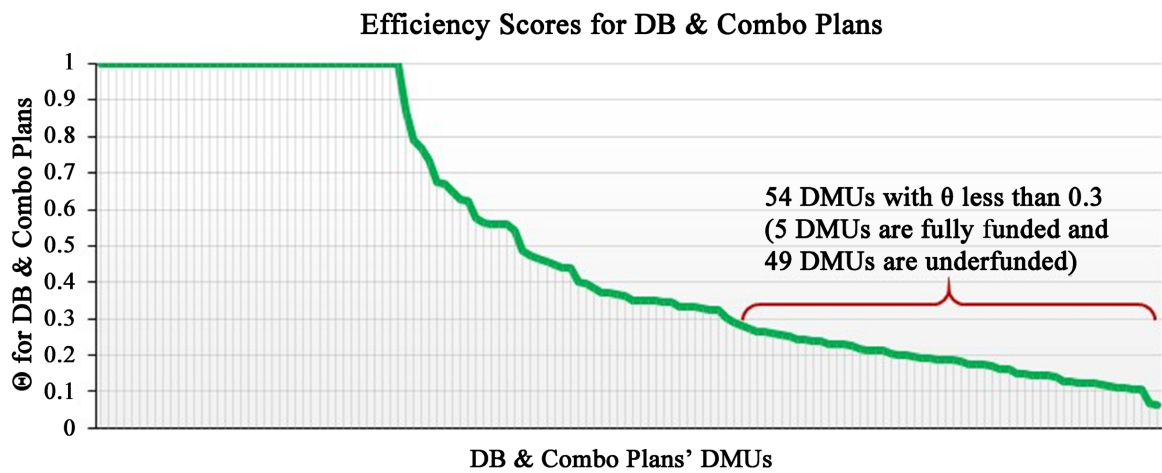## Efficiency Scores for DB & Combo Plans

Figure 2. Efficiency scores' distribution for DB and combo plans.

a) In the first step, the efficient DMUs are extracted from the data. Out of 90 DB plans 33 DMUs are efficient (23 fully funded DMUs and 10 underfunded DMUs)[4]. As presented in **Table 3** in Section 3.2, the average efficiency scores and minimum efficiency score for the output oriented Non-Dis-VRS model considering funds' status for all 90 DB plans are 0.576 and 0.106 respectively. After removing the first layer of efficient DMUs, 57 DMUs remain that the average efficiency scores and minimum efficiency scores for the Non-Dis-VRS model with and without considering categorical situation[5] are presented in **Table 5**:

---

[4]It should be noted that all pension plans for this research are active plans. Therefore, if an active plan is flagged as underfunded it means that the plan will not meet its obligations based on OSFI tiers' deadlines in the next 5 or 10 years. As a result, an underfunded active plan for the year of study for this research can be placed on the frontier as an efficient plan. However, only a few underfunded active plans become efficient.

[5]Considering funds' status (fully funded/underfunded) provides a more realistic analysis of the results. By having categorical DMUs, fully funded plans are referenced only to their own category while the underfunded plans are referenced to their level as well as the higher level of hierarchy (fully funded plans). Therefore, the main change between the Non-Dis-VRS model with and without categorical DMUs is in the DMUs' reference sets. The results of the DEA model with categorical DMUs change slightly and one of the important aspects of pension funds is taken into consideration.

Table 5. Removing first layer of efficient DB plans.

DB Plans # DMUs: 57	Non-Dis-VRS_O_CAT[a]	Average: 0.775 Min: 0.305
	Non-Dis-VRS_O[b]	Average: 0.767 Min: 0.305

[a]Output Oriented Non-Dis-VRS model considering categorical DMUs (fully funded and underfunded pension plans); [b]Output Oriented Non-Dis-VRS model (without considering funds' status).

The results indicate that most of the efficient DMUs are from fully funded pension plans. As expected, when the efficient DMUs are extracted from the dataset, the efficiency scores of the highly inefficient DMUs (which are from underfunded plans) increase significantly.

b) In the second step, out of 57 DB plans, 25 DMUs are efficient (6 fully funded DMUs and 19 underfunded DMUs). After removing these 25 DMUs from the dataset (32 DMUs remain), there are no more fully funded DB plans in the dataset. The results are shown in Table 6.

Table 6. Removing the second layer of efficient DB plans.

DB Plans # DMUs: 32	Non-Dis-VRS_O_CAT	N/A[a]
	Non-Dis-VRS_O	Average: 0.843 Min: 0.414

[a]There are no more fully funded plans in the dataset. Therefore, we do not have categorical DMUs in the model.

The process is stopped here since there are no more fully funded plans in the dataset. Most of the fully funded plans are efficient in the first step and the rest become efficient in the second step.

2) DB and Combo Plans

There are 136 DMUs for DB and Combo plans in which there are 37 fully funded plans and 99 underfunded plans:

a) In the first step for DB and Combo plans, out of 136 DMUs, 39 DMUs are efficient of which 27 DMUs are fully funded plans (23 DB and 4 Combo plans) and 12 DMUs are underfunded (10 DB and 2 Combo plans). As indicated in Table 3 in Section 3.2, the average efficiency scores and minimum efficiency score for DB and Combo pension plans are 0.508 and 0.064 respectively. The results are shown in Table 7 after removing the first layer of efficient DMUs (39

Table 7. Removing first layer of efficient DB and combo plans.

DB & Combo Plans # DMUs: 97	Non-Dis-VRS_O_CAT	Average: 0.738 Min: 0.228
	Non-Dis-VRS_O	Average: 0.71 Min: 0.211

DMUs) from the dataset (97 DMUs remain).

b) In the second step, out of 97 DMUs from part a, 39 DMUs are efficient of which 11 DMUs are fully funded and 28 DMUs are underfunded. The results, after removing these 39 DMUs from the dataset (58 DMUs remain), are presented in Table 8.

Table 8. Removing second layer of efficient DB and combo plans.

DB & Combo Plans # DMUs: 58	Non-Dis-VRS_O_CAT	Average: 0.763 Min: 0.277
	Non-Dis-VRS_O	Average: 0.713 Min: 0.277

c) In the third step, out of 58 DMUs, 26 DMUs are efficient of which 5 DMUs are fully funded and 21 DMUs are underfunded. There are no more fully funded plans in the dataset after this step. The results, after removing these 26 DMUs from the dataset (32 DMUs remain), are reported in Table 9.

Table 9. Removing third layer of efficient DB and combo plans.

DB & Combo Plans # DMUs: 32	Non-Dis-VRS_O_CAT	N/A
	Non-Dis-VRS_O	Average: 0.902 Min: 0.473

The experimental analysis is stopped here since there are no more fully funded plans in the dataset and the results indicate that all fully funded DB and Combo plans become efficient in three steps.

As expected, this approach shows that fully funded plans perform better than underfunded plans. Also, as the number of such plans decrease with each stripping process the remaining plans are more and more underfunded until all are underfunded, the minimum efficiency scores increase significantly.

3.3.2. Stripping DMUs with Low θ

In this section, the aim is to demonstrate how many of the inefficient DMUs with low θ are from underfunded plans and if the test is started from the bottom of the efficiency scores' ranking, after how many underfunded plans the next inefficient DMU is from fully funded plans.

1) DB Plans

When the DB plans are investigated from the bottom of the efficiency scores' ranking, the first inefficient fully funded plan is reached after removing 29 inefficient underfunded DMUs. As expected, this is another indication that fully funded plans perform better than underfunded plans and the reason for having some low θs is because both fully funded and underfunded plans are considered in the DEA models.

The results, after removing these 29 DMUs from the dataset (61 DMUs remain), are reported in Table 10.

Table 10. Removing inefficient underfunded DB plans from the bottom of the efficiency scores' ranking.

DB Plans # DMUs: 61	Non-Dis-VRS_O_CAT	Average: 0.801 Min: 0.328
	Non-Dis-VRS_O	Average: 0.789 Min: 0.282

In summary, the first inefficient fully funded plan is reached after removing 29 inefficient underfunded DMUs from the bottom of the efficiency scores' ranking and the minimum efficiency score after removing these 29 DMUs increased significantly.

2) DB and Combo Plans

After removing the underfunded DB DMUs and reaching the first inefficient fully funded DB plan from the bottom of efficiency scores' ranking as mentioned in Section 3.3.2.1, the Non-Dis-VRS models are run for a mix of DB (61 plans) and Combo plans (46 plans) with 107 DMUs. There are 21 inefficient DMUs from the bottom of efficiency scores' ranking which are all from underfunded Combo plans before reaching the first inefficient fully funded pension plan. In this part, the inefficient underfunded DMUs from the bottom of efficiency scores' ranking are removed from the dataset. The results, after removing these 21 DMUs from the 107 DMUs in the dataset (86 DMUs remain), are shown in Table 11.

Table 11. Removing inefficient underfunded db and combo plans from the bottom of efficiency scores' ranking.

DB & Combo Plans # DMUs: 86	Non-Dis-VRS_O_CAT	Average: 0.742 Min: 0.301
	Non-Dis-VRS_O	Average: 0.732 Min: 0.282

In summary, working from below, the first inefficient fully funded DMU is reached after removing 21 inefficient underfunded Combo DMUs. Note that 29 inefficient underfunded DB plans have already been removed.

As anticipated, this approach demonstrates from another perspective that most of underfunded plans have low efficiency scores.

3.3.3. Investigating Fully Funded and Underfunded Pension Plans Separately

1) DB Plans

In this section, each type of funds' status is examined separately to investigate and visualize the average efficiency scores and minimum efficiency scores for each category of fully funded and underfunded DB plans. The results are presented in Table 12.

2) DB and Combo Plans

Table 12. Examining fully funded and underfunded DB plans separately.

DB Plans Non-Dis-VRS_O[a]	Fully Funded DB Plans #DMUs: 29	Average: 0.883 Min: 0.304
	Underfunded DB Plans #DMUs: 61	Average: 0.66 Min: 0.172

[a]Since fully funded and underfunded pension plans are examined separately, the categorical DMUs do not need to be considered.

The results for each category of fully funded and underfunded DB and Combo plans are presented in Table 13.

Table 13. Examining fully funded and underfunded DB and combo plans separately.

DB & Combo Non-Dis-VRS_O	Fully Funded DB & Combo Plans #DMUs: 37	Average: 0.759 Min: 0.183
	Underfunded DB & Combo Plans #DMUs: 99	Average: 0.643 Min: 0.172

The results show that when fully funded and underfunded plans are tested separately, the average and minimum efficiency scores are higher than when both fully funded and underfunded plans are added together.

In summary, all these three approaches indicate that when fully funded and underfunded plans are examined together the minimum efficiency score decreases significantly. Since most plans are underfunded and both fully funded and underfunded plans should be considered together, very low minimum efficiency scores are often found in this industry.

4. Conclusion

Government rules and restrictions are one of the main issues in the pension funds industry impacting asset allocation and managers' control. This study quantified the impact of the regulations on asset allocation and incorporated the controllability of the variables from the managers' perspective. As such, this research has successfully integrated the main characteristics of pension funds into the DEA model leading to a more realistic assessment of the performance of this financial investment vehicle. Also, one of the issues in the pension funds industry is that most plans are underfunded. To have a comprehensive study of this industry both fully funded and underfunded plans should be considered. An analysis was carried out to assess the fully funded and underfunded plans' positions on the efficient frontier. The results indicated that most fully funded plans are efficient and after two or at most three times of removing the efficient DMUs from the dataset, only underfunded plans remain. As expected, the minimum efficiency score increased significantly after the removal of each efficient frontier. Also, most underfunded plans were located at the bottom of the efficiency scores' ranking. When fully funded and underfunded pension plans were evaluated sepa-

rately, the minimum efficiency score increased compared to when they were considered together. Therefore, a low minimum efficiency score is often found for the pension funds industry since both fully funded and underfunded pension plans should be considered in the model.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Asmild, M., Paradi, J., Aggarwall, V., & Schaffnit, C. (2004). Combining DEA Window Analysis with the Malmquist Index Approach in a Study of the Canadian Banking Industry. *Journal of Productivity Analysis, 21,* 67-89. https://doi.org/10.1023/B:PROD.0000012453.91326.ec

Badrizadeh, M. (2017). *Evaluating Pension Funds Considering Unobservable Variables & Bridging Pension Funds & Mutual Funds through the Development of a New DEA Model.* Ph.D. Dissertation, University of Toronto.

Banker, R. D., & Morey, R. (1986). Efficiency Analysis for Exogenously Fixed Inputs and Outputs. *Operation Research, 34,* 513-521. https://doi.org/10.1287/opre.34.4.513

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science, 30,* 1078-1092. https://doi.org/10.1287/mnsc.30.9.1078

Barrientos, A., & Boussofiane, A. (2005). How Efficient Are Pension Fund Managers in Chile? *Revista de Economia Contemporânea, 9,* 289-311. https://doi.org/10.1590/S1415-98482005000200003

Barros, C. P., & Garcia, M. T. M. (2006). Performance Evaluation of Pension Funds Management Companies with Data Envelopment Analysis. *Risk Management and Insurance Review, 9,* 165-188. https://doi.org/10.1111/j.1540-6296.2006.00092.x

Barros, C. P., & Peypoch, N. (2009). An Evaluation of European Airlines' Operational Performance. *International Journal of Production Economics, 122,* 525-533. https://doi.org/10.1016/j.ijpe.2009.04.016

Basso, A., & Funari, S. (2005). A Generalized Performance Attribution Technique for Mutual Funds. *Central European Journal of Operations Research, 13,* 65-84.

Blancard, S., & Martin, E. (2013). Energy Efficiency Measurement in Agriculture with Imprecise Energy Content Information. *Energy Policy, 66,* 198-208. https://doi.org/10.1016/j.enpol.2013.10.071

Charles, V., Aparicio, J., & Zhu, J. (2020). *Data Science and Productivity Analytics* (1st ed.). Springer. https://doi.org/10.1007/978-3-030-43384-0

Charnes, A., Cooper, W. W., & Rhodes E. (1978). Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research, 2,* 429-444. https://doi.org/10.1016/0377-2217(78)90138-8

Chiu, Y. H., Chyanlong, J., Shen, D. B., & Wang, P. C. (2008). Efficiency and Capital Adequacy in Taiwan Banking: BCC and Super-DEA Estimation. *The Service Industires Journal, 28,* 479-496. https://doi.org/10.1080/02642060801917604

Cooper, W. W., Seiford, L. M. and Tone, K. (2007). Data Envelopment Analysis, A Comprehensive *Text with Models, Applications, References and DEA-Solver Software* (2nd ed.). Springer Science and Business Media, LLC.

Cooper, W. W., Seiford, L. M., & Zhu, J. (2011). *Handbook on Data Envelopment Analysis* (2nd ed.). Springer Science and Business Media, LLC.
https://doi.org/10.1007/978-1-4419-6151-8

Field, A. (2009). *Discovering Statistics Using SPSS* (3rd ed). SAGE Publications Inc.

Financial Services Commission of Ontario (2004). *A Proposed Risk-Based Investment Monitoring Model for Supervising Defined Benefit Pension Plans. Consultation Paper.*
http://www.fsco.gov.on.ca/en/Pages/default.aspx

Galagedera, D. U. A., & Watson, J. (2015). Benchmarking Superannuation Funds Based on Relative Performance. *Applied Economics, 47,* 2959-2973.
https://doi.org/10.1080/00036846.2015.1011315

Garcia, M. T. M. (2010). Efficiency Evaluation of the Portuguese Pension Funds Management Companies. *Journal of International Financial Markets, Institutions & Money, 20,* 259-266. https://doi.org/10.1016/j.intfin.2010.03.003

Paradi, J. C., Sherman, H. D., & Tam, F. K. (2018). *Data Envelopment Analysis in the Financial Service Industry. A Guide for Practitioners and Analysts Working in Operations Research Using DEA* (1st ed.). Springer.
https://doi.org/10.1007/978-3-319-69725-3

Premachandra, I. M., Zhu, J., Watson, J., & Galagedera, D. U. A. (2012). Best Performing US Mutual Fund Families from 1993 to 2008: Evidence from a Novel Two-Stage DEA Model for Efficiency Decomposition. *Journal of Banking and Finance, 36,* 3302-3317.
https://doi.org/10.1016/j.jbankfin.2012.07.018

Sathye, M. (2011). The Impact of Financial Crisis on the Efficiency of Superannuation Funds. *Journal of Law & Finance Management, 10,* 16-27.

Simar, L. (2003). Detecting Outliers in Frontier Models: A Simple Approach. *Journal of Productivity Analysis, 20,* 391-424. https://doi.org/10.1023/A:1027308001925

Smith, P. (1997). Model Misspecification in Data Envelopment Analysis. *Annals of Operation Research, 73,* 233-252. https://doi.org/10.1023/A:1018981212364

Vlontzos, G., Niavis, S., & Manos, B. (2014). A DEA Approach for Estimating the Agricultural Energy and Environmental Efficiency of EU Countries. *Renewable and Sustainable Energy Reviews, 40,* 91-96. https://doi.org/10.1016/j.rser.2014.07.153

Wanke, P., & Barros, C. P. (2016). Efficiency in Latin American Airlines: A Two-Stage Approach Combining Virtual Frontier Dynamic DEA and Simplex Regression. *Journal of Air Transport Management, 54,* 93-103. https://doi.org/10.1016/j.jairtraman.2016.04.001

Zamuee, M. R. (2015). Data Envelopment Analysis to Measure Efficiency of Namibian Pension Funds. *American Journal of Marketing Research, 1,* 215-221.