

# Comparing the Efficiency of Stores at New Zealand Post

Harriet Priddey and Kane Harton  
Department of Engineering Science  
University of Auckland  
New Zealand  
hpriiddey@gmail.com, kharton@gmail.com

---

## Abstract

New Zealand Post operates more than one hundred Post Stores around the country. They wished to find a way of comparing these stores impartially. The method used to do so in this work is Data Envelopment Analysis (DEA), a performance measurement technique utilising linear programming. Each store uses inputs to produce outputs, and an efficiency score is calculated based on how well it does so.

A significant part of the work involved selecting the inputs and outputs for the model, done using discussion, sensitivity analysis and statistical tests. To solve the DEA model, software was developed using the Python language and PuLP module. Extensions to the basic model included non-discretionary variables, weight restrictions, virtual weights and a second phase LP.

Next, the results of the DEA model were analysed. This involved identifying efficient stores and areas where they provide an example of good practice. Also, inefficient stores were assigned efficient peers to emulate, and efficiency targets. Some innovative ways of investigating the data included graphs, virtual weights and totalising monetary factors.

**Key Words:** Data Envelopment Analysis, Efficiency, Post Office

---

## 1 Introduction

New Zealand Post is one of the largest and most visible businesses in the country. They own and operate many retail stores around the country. They wished to find a way to compare these stores fairly and identify good and poor performers.

The approach used to do so in this work is Data Envelopment Analysis. This technique involves formulating and solving linear programs. Each store receives an efficiency rating based on how well it converts inputs to outputs, compared to other stores. This rating allows each store to choose its input and output weights so that the weighted input/output ratio is as good as possible. If, even with these weights, there is another store with a better ratio then the first store is inefficient. The results of the analysis also provide efficient peers for inefficient stores, and allow us to identify areas where stores perform particularly well.

The most appropriate inputs and outputs for the DEA model were identified using discussion, sensitivity analysis and statistical tests. The results of the model were reported to New Zealand Post in tabular and graphical form. They were analysed further

by looking at input and output weights, and carrying out additional analyses using totalised values. To do this, the solver DEA.py was developed using Python and PuLP.

## 2 DEA Theory

Data Envelopment Analysis, commonly abbreviated to DEA, is a non-parametric performance measurement technique. It compares Decision Making Units (DMUs) in a multiple input, multiple output situation. DMUs are the organisation entities concerned with making decisions regarding inputs and outputs. A DMU is considered fully efficient “if and only if the performances of other DMUs does not show that some of its inputs or outputs can be improved without worsening some of its other inputs or outputs” (Cooper, Seiford, & Zhu, 2004).

There are many different DEA models. The one used in this work was proposed by Banker, Charnes and Cooper (1984) and is commonly known as the BCC model. A mathematical program is formulated and solved for each DMU. This program aims to maximise the ratio of weighted outputs to weighted inputs, known as the efficiency, subject to the efficiency of all DMUs being less than or equal to one. The model shown here is a linear form that eliminates the fractional ratio. It takes into account variable returns to scale. This means that differently sized DMUs may have different proportional changes in outputs when inputs change.

There are two forms of the BCC model, the envelopment form and the multiplier form. These two forms satisfy the primal-dual relationship of linear programming, thus the same optimal objective (the efficiency score) will be found by solving either form. The equations are detailed below, using notation from Cooper, Seiford and Zhu (2004).

### Indices

- $i$  = input:  $1, \dots, m$ .
- $r$  = output:  $1, \dots, s$ .
- $o$  = DMU that is being solved for.
- $j$  = DMU:  $1, \dots, n$ .

### Parameters

- $x_{ij}$  = amount of input consumed by DMU  $j$ .
- $y_{rj}$  = amount of output produced by DMU  $j$ .

### Decision variables

- $\theta = z$  = efficiency score
- $\lambda_j$  = weighting of DMU  $j$  for current DMU.
- $s_i^-, s_r^+$  = slack for input  $i$  or output  $r$ .
- $u_r$  = weighting of output  $r$  for current DMU.
- $v_i$  = weighting of input  $i$  for current DMU.
- $\mu_o$  = dual variable, relates to returns-to-scale.

### Model E: BCC Input-oriented Envelopment Form

$$\begin{aligned} & \text{Minimise } \theta - \varepsilon(\sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+) \\ \text{(E1)} \quad & \sum_{j=1}^n x_{ij} \lambda_j + s_i^- = \theta x_{io} && \text{for } i = 1, 2, \dots, m. \\ \text{(E2)} \quad & \sum_{j=1}^n y_{rj} \lambda_j - s_r^+ = y_{ro} && \text{for } r = 1, 2, \dots, s. \\ \text{(E3)} \quad & \sum_{j=1}^n \lambda_j = 1. \\ & \lambda_j \geq 0. \end{aligned}$$

**Model M: BCC Input-oriented Multiplier Form**

$$\begin{aligned}
 & \text{Maximise } z = \sum_{r=1}^s \mu_r y_{r0} - \mu_o \\
 \text{(M1)} \quad & \sum_{r=1}^s \mu_r y_{rj} - \sum_{i=1}^m v_i x_{ij} - \mu_o \leq 0 \quad \text{for } j = 1, 2, \dots, n. \\
 \text{(M2)} \quad & \sum_{j=1}^n v_j x_{j0} = 1. \\
 & u_j, v_i \geq 0, \mu_o \text{ free in sign.}
 \end{aligned}$$

### 3 The New Zealand Post Problem

The New Zealand Post group is the second largest retailer in New Zealand, with over 17,000 employees. There are over 300 Post Shops and 650 Post Centres around the country (New Zealand Post, 2010). Services available include postal solutions, bill payments, Kiwibank services, car registration, Bonus Bonds, travel bookings and insurance. These retail stores operate in varying conditions around the country, from major urban cities to rural areas. Stores have differing levels of competition within the area, serve populations of different socio-economic backgrounds, and have different costs, profits and levels of customer satisfaction. The stores are split into three categories (T1, T2, and T3) depending on the banking functions they provide.

With all these factors to consider, New Zealand Post wished to find a way of comparing stores fairly, and identifying the best and worst performers. In particular, they wished to look at Post Shops, which provide more services than Post Centres. So they approached the Department of Engineering Science. The solution proposed was to use DEA, which would allow stores to be compared based on a variety of factors. This would also allow each store to choose which factors to be assessed on, rather than using pre-determined weightings.

## 4 Model Development

### 4.1 Process for Choosing Inputs and Outputs

In this section we discuss the process of deriving inputs and outputs to measure efficiency using DEA. We highlight different types of inputs and outputs that needed special treatment within the DEA analysis and also mention difficulties encountered with some of them.

Identifying the inputs and outputs to use in a DEA assessment is crucial, and arguably the most important part of the whole process (Thanassoulis, 2001). Choosing different inputs and outputs can significantly affect the results, and make different DMUs appear better or worse. It is important to use as few factors as possible, because too many factors means that many DMUs can appear efficient by choosing their own particular input and output weightings. This reduces the discriminating power of the DEA analysis.

It is also very important to get input from the decision-makers when choosing inputs and outputs. They know the practical realities, and may have crucial insight that is unavailable to the analyst. If decision makers are not consulted then they may later contest the results, reducing the value of the DEA assessment. In this work, New Zealand Post analysts and store managers were involved through the process.

There were several meetings with New Zealand Post for this work. A list of initial inputs and outputs was compiled, then narrowed down based on relevance and which data was actually available. This provided the basis for the initial DEA model. Later there was another meeting with New Zealand Post, including several store managers. A large number of inputs and outputs were suggested. These were then refined based on

what was available, further discussion, statistical tests and sensitivity analysis. Some specific issues with choosing inputs and outputs are detailed in the following sections.

#### **4.2 Types of Post Stores**

New Zealand Post divides its stores into three types, depending on the services they offer. These types relate to whether a store has the ability to grant home loans and perform some other banking functions. This increases a store's ability to generate revenue. Some stores have full banking functions (T3), some have limited functions (T2), and some have none (T1).

The initial model carried out separate DEA analyses for T1 stores and T3 stores. We raised the possibility of including both types of stores in a single model and using categorical variables. However New Zealand Post management believes the stores are far too different to include together. In DEA, if the categories are not comparable then a separate analysis should be performed for each category (Cooper, et al., 2004).

#### **4.3 Non-discretionary Inputs**

Non-discretionary inputs are those which influence performance, but which the given DMU has little or no influence over. There exist factors which post store management cannot control, but which may influence performance, for example store floor area or median income in an area. Hence, the model should take this into account.

This was done using the formulation proposed by Banker and Morey (1986). This model splits the inputs into two subsets, discretionary ( $I_D$ ) and non-discretionary ( $I_N$ ). The non-discretionary input constraint does not include  $\theta$  on its right-hand side, meaning these inputs cannot be reduced. This model was coded into the DEA Python program.

#### **4.4 Competition Input**

Competition was considered to be an important input as it affects how well a store can perform. Data was collected in order to get a competition 'score' for each store. The physical address was found, and then Google Maps was used to identify the latitude and longitude (Harton, 2010a). The number of other post stores within 5km was found using an Excel spreadsheet and a macro. In DEA, a store is more efficient if it uses fewer resources. So, the number of post stores within 5km was treated as an undesirable input (Harton, 2010b). A number was applied to each store, lower if there were more post stores within the area, higher if there were fewer. This way a post store in a competitive environment will appear more efficient if it performs the same as another store in an 'easier' environment.

#### **4.5 Staff Engagement**

One of these issues related to a variable in the initial DEA model called engagement, which represented staff satisfaction with their working conditions. There were a number of difficulties with including this variable. For one, its measurement could be unreliable and dependent on a staff member's mood on one particular day. Also, it was not clear whether engagement should be included as an input or an output. It could be considered as an input because staff members who are more engaged contribute to earning more revenue and producing better outputs. However by doing this, the DEA model rewarded stores who minimised the input while still achieving outputs, i.e. had very unsatisfied staff. Engagement could also be included as an output because it could be considered an objective to have contented staff.

Engagement was trialled in the model, but because of these problems it was removed. It had some effect on results, as fewer stores were efficient when it was not included. However because of the interpretation problems, it did not make sense to include it.

#### **4.6 FTE and Staff Costs**

In the original data we had two related inputs, FTE (Full-time equivalent, the number of staff working at the store) and Staff Costs. It was expected that these inputs would be correlated, as employing more staff (higher FTE) will generally lead to higher staff costs. So a regression analysis was done which found that the  $R^2$  values for correlation between Staff Costs and FTE were 65.55% for T1 and 51.79% for T3 stores. The T1 and T3 categorisations relate to the banking functions of a Post Store. The closer an  $R^2$  value is to 1, the stronger the correlation, so this suggests that the variables are correlated, though not excessively.

Another way to investigate the correlation was to try a sensitivity analysis, by running the DEA twice more, with only one of Staff Costs and FTE included each time. This showed that fewer stores were considered efficient when only one of the inputs is used, as opposed to both. This is as would be expected.

Overall, both inputs could be left in the model to allow for different scenarios. An example of this is many long term staff members who have resulting higher pay, versus part time or temporary staff. Including both inputs does not result in many more stores becoming efficient. Usually in DEA it is not a bad thing to include correlated factors, as it allows for slightly different situations to be accommodated. If the factors are fully correlated then the same stores would be efficient whichever factor was included. However, this must be balanced against the possibility of including too many inputs and outputs so that results become meaningless.

#### **4.7 Mystery Shopper Results Output**

Mystery shopper results were a potential output for the DEA model. Mystery shoppers are customers sent into stores on certain days to measure the level of service. They report back using a questionnaire, and the scores are averaged in a year-to-date (YTD) value. There was debate over whether to include this. It is more variable than the expenditure-based measures, as it depends on which day it is measured. New Zealand Post did not want stores appearing efficient only through their mystery shopper rating.

To deal with this we used two approaches. The first was to simply exclude mystery shopper results, and then do a sensitivity analysis to examine the effects. This resulted in 44 efficient stores without including mystery shopper results, compared to 57 with it included, for the T1 data. This is desirable, as 57 efficient stores out of 91 in total are too many to provide valuable information. However, mystery shopper results represent customer satisfaction, and as such it would be good to include it in the model in some form.

The second approach was to use a weight restriction. There are several different ways to restrict weights, including absolute, relative and input-output weight restrictions (Cooper, et al., 2004). The most appropriate type in this situation was virtual weight restrictions. Virtual weights are the input/output weight multiplied by the actual value of the input/output, and they loosely represent the proportion of the efficiency score that is determined by the particular input or output. These were used rather than actual weights because of the large difference in scale between mystery shopper results and the other

variables, meaning mystery shopper weights were disproportionately large. It was decided that mystery shopper should be restricted to a virtual weight of below 0.3. This loosely corresponds to contributing no more than 30% towards the efficiency score.

When comparing the results it was found that adding a weight restriction reduced the number of efficient stores. For the T1/T2 data, 50 stores were considered efficient with the weight restriction compared to 56 without. There were also several inefficient stores which had a lower efficiency score after the weight restriction. This supports the conclusion that weight restrictions allow more discrimination within the model.

#### **4.8 Location**

Location was provided as a categorical variable. Each store was in one of three categories, based on Statistics New Zealand definitions – main urban area, satellite urban area or independent urban area (Statistics New Zealand, 2010). Main urban areas are situated in major cities and towns. Satellite urban areas are settlements with strong links to a main urban area, while independent urban areas do not have such links; they are in more rural areas. There are a number of ways to include categorical variables in a DEA model, the most common of which was outlined by Banker and Morey (1986). It involves ranking the categories in order of most to least difficult to operate in. Then each DMU is compared to other DMUs who operate in the same or more difficult conditions. This means DMUs operating in difficult conditions are not unfairly penalised by being compared to DMUs who have it easier.

However, in the New Zealand Post situation it was not clear how the categories should be ordered. There are arguments for each way. Main urban area stores have a larger population to draw customers from, however independent urban area stores have less competition, and tend to have a steadier workforce and customer base.

So an initial DEA analysis was carried out without location included. The results from this were grouped according to which location category they came from. It was found for T1 data that 84% of independent urban area stores were efficient compared to 58% of main urban area stores. For T3 data, 100% of independent urban area stores were efficient compared to 51% of main urban area stores. This suggests that rural conditions can be considered easier to operate in. There were not enough satellite urban areas stores to gain useful results, however we can surmise that their difficulty would be between that of independent urban area and main urban area.

## 5 Final Model

The inputs and outputs of the final DEA model are shown in figure 1.

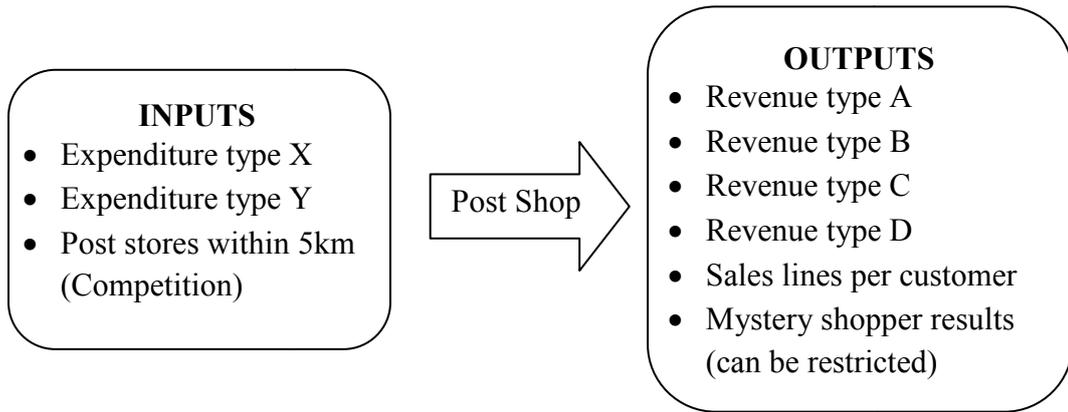


Figure 1: The Final New Zealand Post DEA Model

## 6 Open Source Python DEA Software

### 6.1 DEA Software

A program to solve DEA problems was written and developed. This program is called DEA.py, and is written in the Python language (Python Software Foundation, 2010). It uses the python library PuLP (PuLP Documentation Team, 2010) to write a series of linear programs which are then solved using the free solver coinOR (COIN-OR, 2010). It also uses the Python packages xlrd and xlwt for reading and writing to Microsoft Excel and yapgvb for producing a graph. All the components are free and open source. Some particular features of DEA.py are detailed in the following sections.

The input to this program is an Excel file, which needs to have specific formatting. The first column is a list of the DMUs, starting in the second row. There is a gap of one column, then a block of the inputs. The name of the input is in the top row, then below that the input values for each DMU. Next is another column gap then the non-discretionary inputs, if required. Finally comes a block of the outputs.

DEA.py can be run by opening it in Notepad or a similar text editor and change the required variables, then save. Then open a command window and enter the line `[location of python] python [location of DEA.py] \DEA.py`. DEA.py generally solves the New Zealand Post problem in well under a minute. The more DMUs in the problem, the longer it will take, because an LP is solved for each DMU.

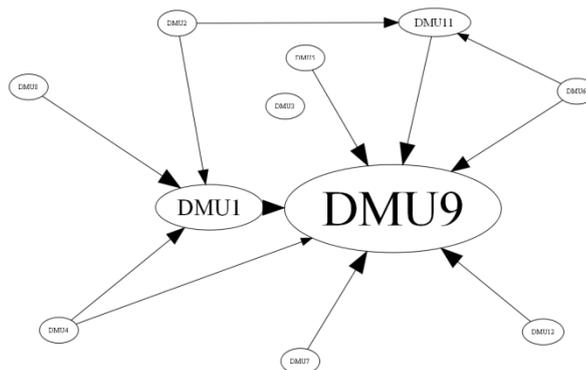


Figure 2: Graph produced by DEA.py for a small example problem

The Excel file output contains the solution to the DEA model. It gives the efficiency scores for each DMU, efficient peers, benchmarked DMUs, weights, targets and slacks.

An example of the output graph produced is shown in figure 2. The output graph represents each DMU as an oval. Arrows point from an inefficient DMU to its efficient peers, with larger arrow heads showing a larger weighting. The oval for a DMU is larger the more times it is referenced. Some DMUs are not linked to any others, meaning they are efficient but do not act as a peer to any others. This indicates that they may have somewhat different circumstances to any other DMUs. This graph gives an impression of which DMUs are most useful as examples of good practice.

## **6.2 Virtual Weights**

Several extensions were added to the basic DEA software. One was adding a sheet that output the weighted data, also known as virtual weights. This is the input/output weight multiplied by the actual value of the input/output. Virtual weights do not depend on units of measurement (Cooper, et al., 2004), which is important in this case because inputs and outputs have very different scales. This allows comparison of the importance of inputs and outputs, and in addition can be easily graphed for visual comparison.

## **6.3 Envelopment and Multiplier Form**

There are two versions of DEA.py, one solves the envelopment form of the model, the other solves the multiplier form. This allows results to be checked, and may provide alternative set of input/output weights. Solving DEA problems with the envelopment form is generally be faster to solve as there are fewer constraints. Using the multiplier model makes it easier to implement weight restrictions.

## **6.4 Weight Restrictions**

Generally in DEA, all DMUs are free to choose their own input and output weights, in order to maximise efficiency, while maintaining feasibility for all other DMUs (Cooper, et al., 2004). An extension of the model was to add in input/output weight restrictions. This is desirable in cases where we know that some factors are not as important as others, and thus efficiency scores should not be overly influenced by these factors.

Absolute weight restrictions simply limit weights to within a specific range (Cooper, et al., 2004). These can be coded into DEA.py by following the instructions given in the code. A future development would be to allow weights restrictions to be included in the input sheet instead of hard coding.

## **6.5 Second Phase LP**

Another addition was to implement a second phase LP. A DMU should not be considered efficient if it ignores an input or output efficiently. However this means that the non-Archimedean  $\epsilon$  is included in the objective of the full DEA formulation. This is not possible to code. So the second stage LP attempts to maximise the sum of slacks. A DMU with an efficiency score of 100% is only truly efficient if its maximum sum of slacks is zero, meaning it is properly on the efficient frontier.

It was found that weakly efficient DMUs were not an issue for the New Zealand Post situation, as all stores with efficiency of 100% had their maximum sum of slacks equal to zero. This means that the efficiency score can safely be considered by itself. This was somewhat expected, as for the second phase to be relevant there would need to be DMUs that have the same values for a number of inputs and outputs. However if this

software were to be used on a different data set, it may be very useful to have the second phase problem implemented.

## 7 Analysis

The results for this model, obtained from DEA.py, have been handed to New Zealand Post; however they cannot be given in full in this report because of privacy issues. Instead some examples of how the data was analysed are given.

The efficiency scores are the first useful information. Stores with a score of 100% are fully efficient. Others are inefficient, and should theoretically be able to improve their performance to reach the efficient frontier. It is also interesting to look at the efficient peers for each inefficient store. They provide a group of stores with similar strengths who can be used as an example of good practice. Reversing this, we can also look at the number of times a store acts as an efficient peer, or is ‘benchmarked’.

The virtual weights are of interest, particularly for efficient stores. They indicate which areas a store performs particularly well in. It may be useful to investigate why this is. It should however be remembered that there can be multiple optimal solutions for the weights. For inefficient stores it is useful to look at the targets. These show how much a store should increase outputs or decrease inputs to become efficient.

Some graphs were also created to present the results to New Zealand Post, such as a bar graph of the percentage targets (figure 3) and a pie graph of the weights for efficient peers. It was also interesting to total up revenue and expenditure values so that a two-dimensional DEA analysis could be carried out, for easier graphing.

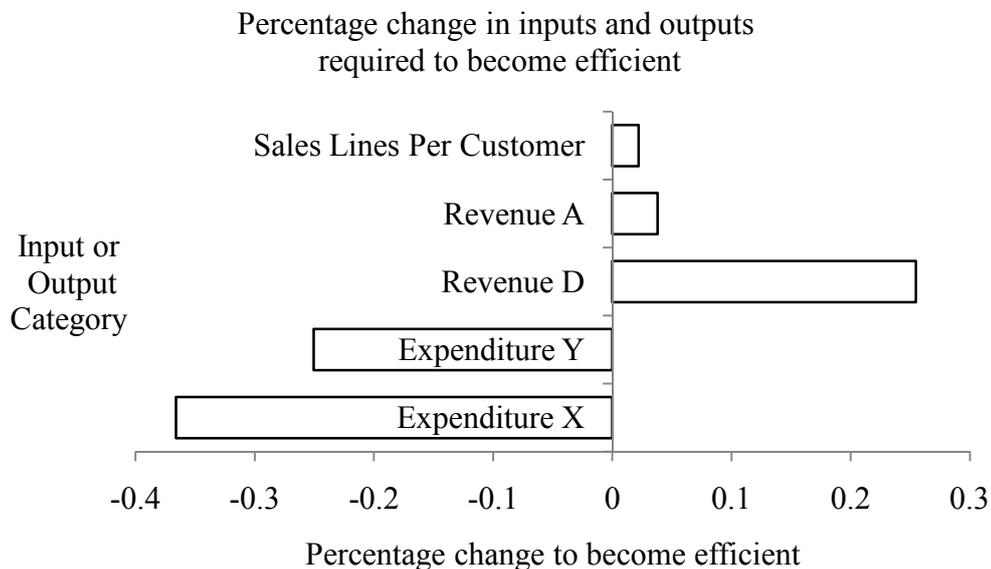


Figure 3: Percentage targets for a particular inefficient store

## 8 Conclusions

This project has applied Data Envelopment Analysis to New Zealand Post’s retail stores. DEA is not a one-size-fits-all process; instead it must be applied differently depending on the situation, available data, and needs of the decision-makers. Choosing inputs and outputs is a vital part of the DEA model, and judgement is required. In this work statistical analysis, sensitivity analysis and discussion were used to narrow down the potential inputs and outputs.

This work involved the writing and development of the Python program DEA.py. It is now a useful, user-friendly piece of software that incorporates several DEA models and techniques. The results were analysed and examined in several ways.

The DEA process will have a number of benefits to New Zealand Post. They will be able to compare their retail stores, identify the best performers and areas where they do well. They will also identify inefficient stores, efficient peers for them to emulate and targets that they should be able to achieve.

Future work will involve making the solver DEA.py available open source. This means others will be able to develop it further and use it for their own DEA analyses.

## Acknowledgments

Thank you to Dr Andrea Raith and A/Prof Matthias Ehrgott for their suggestions and guidance while supervising this work. Thank you also to Jody Snowdon and Raj Appana from New Zealand Post for their advice and assistance.

## References

- Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, 30(9), 1078-1092.
- Banker, R. D., & Morey, R. C. (1986). The Use of Categorical Variables in Data Envelopment Analysis. *Management Science*, 32(12), 1613-1627.
- COIN-OR. (2010). COmputational INfrastructure for Operations Research. Retrieved November 8, 2010, from <http://www.coin-or.org/>
- Cooper, W. W., Seiford, L. M., & Zhu, J. (2004). *Handbook on data envelopment analysis*. Boston: Kluwer Academic.
- Harton, K. (2010a). *Data Envelopment Analysis with NZ Post*: Department of Engineering Science.
- Harton, K. (2010b). *Using DEA.py*: Department of Engineering Science.
- New Zealand Post. (2010). Our Major Business Streams. Retrieved August 18, 2010, from <http://www.nzpost.co.nz/Cultures/en-NZ/AboutUs/OrganisationalInformation/WhatWeDo/OurMajorBusinessStreams.htm>
- PuLP Documentation Team. (2010). Optimization with PuLP. Retrieved September 20, 2010, from <https://www.coin-or.org/PuLP/>
- Python Software Foundation. (2010). Python Programming Language – Official Website. Retrieved September 1, 2010, from <http://www.python.org/>
- Statistics New Zealand. (2010). Urban/Rural Profile (experimental) Classification Categories. Retrieved September 11, 2010, from [http://www.stats.govt.nz/browse\\_for\\_stats/people\\_and\\_communities/geographic\\_regions/urban-rural-profile-experimental-class-categories.aspx](http://www.stats.govt.nz/browse_for_stats/people_and_communities/geographic_regions/urban-rural-profile-experimental-class-categories.aspx)
- Thanassoulis, E. (2001). *Introduction to the theory and application of data envelopment analysis : a foundation text with integrated software*. Norwell, Mass.: Kluwer Academic Publishers.