Analytics Models for Corporate Social Responsibility in Global Supply Chains

ANALYTICS MODELS FOR CORPORATE SOCIAL RESPONSIBILITY IN GLOBAL SUPPLY CHAINS

BY

SAMEH HABBOUBI, M.Sc.

a thesis

submitted to the School of Computaional Science & Engineering

and the School of Graduate Studies

of McMaster University

in partial fulfilment of the requirements

for the degree of

Master of Science

Master of Science (2018)	McMaster University
(Computational Science and Engineering)	Hamilton, Ontario, Canada

TITLE:	Analytics Models for Corporate Social Responsibility in Global Supply Chains
AUTHOR:	Sameh Habboubi M.Sc., (School of Computational Science & Engineering) McMaster University, Hamilton, Canada
SUPERVISOR:	Dr. Elkafi Hassini
NUMBER OF PAGES:	ix, 85

Abstract

There have been several infamous incidences where world-renowned corporations have been caught by surprise when a low-tier downstream supplier has been publicly found to be non-compliant with basic corporate social responsibilities (CSR) codes. In such instances the company reputation, and consequently financial health, suffer greatly. Motivated by the advances in predictive modeling, we present a predictive analytics model for detecting possible supplier deviations before they become a corporate liability. The model will be built based on publicly available data such as news and online content. We apply text mining and machine learning tools to design a corporate social responsibility " early warning system" on the upstream side of the supply chain. In our literature review we found that there is a lack of studies that focus on the social aspect of sustainability. Our research will help fill this gap by providing performance measures that can be used to build prescriptive analytics models to help in the selection of suppliers. To this end, we use the output of the predictive model to create a supplier selection optimization model that takes into account CSR compliance in global supply chain context. We propose a heuristic to solve the problem and computationally study its effectiveness as well as the impact of introducing CSR on procurement costs as well as ordering and supplier selection patterns. Our models provide analytics tools to companies to detect supplier deviance behaviour and act

upon it so as to contain its impact and possible disruptions that can shake the whole supply chain.

Acknowledgements

It is with a great pleasure that I reserve these few lines of gratitude and deep appreciation to all those who directly or indirectly contributed to the completion of this work.

I express my greatest thanks to my supervisor Dr. Elkafi Hassini who, despite his many duties, supported me with valuable feedback and always kindly encouraged me to succeed in completing my thesis research. His advice, dedication, availability, valuable comments and corrections led to the success of this work.

Finally, I wish to express my deep gratitude and thank my family who has consistently expressed its unconditional support and encouragement. All those who contributed in one way or another to this work, will hopefully find, the crowning of their efforts in this work.

Contents

List of Figures

Chapter 1

Introduction and Problem Statement

1.1 Motivation

Recent literature reviews e.g., (Hassini *et al.*, 2012) and (Tajbakhsh and Hassini, 2015) have concluded that there is a lack of quantitative models that study social sustainability issues in supply chains. The main reason is the lack of objective measures to assess supplier compliance to CSR principles. Our work will help fill in this gap by providing a data-driven method that will provide objective risk scores for suppliers' compliance to CSR principles. Motivated by the advances in machine learning, we propose a predictive analytics model to detect possible supplier deviations before they become a corporate liability. This model uses text mining and classification techniques in order to provide objective probabilistic classifiers that will be the basis of our CSR risk scores. Although in this thesis, we use publicly available data, such as news and on-line content to calculate the CSR risk, our models can easily incorporate companies' proprietary data such as inspections and factory visit reports.

Faced with highly competitive markets, characterized by a high demand for personalized products, with good quality, delivered in minimum time and all at the lowest cost, today's companies are managing their purchases effectively to create a substantial competitive profit. Therefore, supplier selection becomes a strategic decision that will help firms create and maintain a network of reliable and efficient suppliers to meet the growing competition challenges. In the literature, the concept of supplier selection models that help companies manage effectively their purchases and minimize their cost is not new. However, these models often do not include the social sustainability aspect of the supplier as a criteria of selection. We introduce a new supplier selection prescriptive model that incorporates CSR criteria that are built based on the CSR scores from the CSR predictive model. Our model presents a multi-period inventory lot-sizing problem, where there are multiple products and multiple suppliers, with each supplier introducing product-specific capacities. Our model results in a complex mixed integer nonlinear optimization problem. To solve it, we first linearized and then propose a heuristic approaches that produces good quality solutions.

1.2 Problem statement

While the economic and environmental aspects of the sustainability triple bottom line have been generally well covered by academics and practitioners, the social aspect has received much less attention (Tajbakhsh and Hassini, 2015). To address this deficiency, corporate social responsibility (CSR) has evolved from a restricted application of social responsibility to a globally institutionalized function that is integrated with strategic management and corporate governance (Carroll, 2008). Taking CSR to a supply chain level has its challenges given the intricacies in measuring sustainability through a supply chain (Hassini *et al.*, 2012).

The complexities multiply when we consider a global supply chain where products and services may be outsourced from a multi-tiered supply network that extends through different countries. In such complex supply chains it becomes difficult for a firm to have complete visibility of its suppliers and in turn monitor their CSR activities. For example, a company like Boeing makes its 787 plane at over 5000 factories around the world. When its dream-liner had a burning battery problem, the company engineers blamed their suppliers "saying that poor quality components are coming from subcontractors that have operated largely out of Boeing's view" (Gates, 2013). If Boeing engineers have found it difficult to trace and evaluate the quality of a battery supplier, we can only imagine how difficult it will be to evaluate social sustainability practices at such a supplier's premises. It is thus our goal in this thesis to develop a methodology to aid companies in detecting supplier deviance behavior early so as to act upon it and contain its impact and possible disruptions that can shake the whole supply chain. In this regard, we address two important questions:

1. How do we assess the CSR risk of suppliers in an objective manner?

2. How to incorporate the resulting assessment measures in supplier selection models?

To address the first question we adopt a data-driven approach where we can potentially use all company and publicly available data to mine for CSR related information using machine learning and aggregate the findings in CSR scores. These scores are then added to a supplier selection and lot sizing optimization models as attributes of the suppliers' CSR-related risks. We then add a CSR constraint to the model that ensures that a firm will minimize its procurement costs while at the same time meeting its preset CSR goals.

1.3 Contributions

The contributions of the present thesis can be summarized as follows:

1. ***A predictive model based on machine learning for text mining::***
   (a) Building a predictive model to quantify CSR risk probabilities for a given supplier using news articles from the Factiva news database.

2. ***A prescriptive model for supplier selection and lot sizing with CSR considerations:***
   (a) Building a multi-period inventory lot-sizing optimization model that takes the CSR risk scores as input and incorporates a constraint for meeting CSR thresholds.

   (b) Building a procedure to generate feasible problem instances to test and study this optimization model

3. ***Algorithmic:***
   (a) Developing a Heuristic to solve the optimization model

1.4 Thesis structure

The remainder of the thesis is structured as follows. In Chapter 2, we provide a background discussion on the social aspects of sustainability as well as predictive modeling and supplier selection. Our predictive model and its application are discussed in

Chapter 3. Finally, in chapter 4 we present our supplier selection optimization model. We explain how we incorporate the CSR scores from the predictive model in Chapter 3. In addition, we discuss the complexity of the resulting problem and propose a heuristic to solve it. Finally, we report our numerical study findings as well as some managerial insights. We present our thesis conclusions, study limitations, and future work in Chapter 5.

Chapter 2

Literature Review

In this chapter, we present a literature review of the main studies that are related to our thesis. In the first section, we review the concept of corporate social responsibility and its importance in global supply chains. In the second section, we discuss the theory of text mining and predictive modeling. In the third section, we review the literature on supplier selection and lot-sizing models. Finally, we conclude our findings from the literature review and highlight some of the literature gaps and our contributions towards filling those gaps.

2.1 Corporate Social Responsibility

The concept of corporate social responsibility (CSR) describes the integration of social and ecological aspects in entrepreneurial activities as well as the responsibility for the impact of these activities on society. CSR takes in a wide range of social and environmental protection, including working conditions, human rights, environmental protection, prevention of corruption, fair competition, consumer protection and taxes transparency. It is essential that this responsibility fits into the core business of the company. This means that CSR looks at how companies generate their profits and not how they use them in the sense of a "ethics of giving".

Interest in CSR started in the 1980s with the rise of civil society, in a context of growing globalization and at major conferences that resulted in the Brundtland Report (Brundtland, 1987). With this development of environmental, but also social and economic concerns, corporate responsibility is becoming an increasingly important issue. More and more consumers are becoming critical about companies and want them to be more law-abiding, environmentally friendly and more accountable in general.

As a result, governments around the world are putting in place regulations that will lay the foundations for modern CSR. For example, in France, the NRE laws are among the first to oblige firms to report their performance regarding their sustainable development (Egan *et al.*, 2003). Such regulations are forcing companies to invest in CSR, so as not be outdistanced by their competitors and to enhance their reputation in the market (Militaru and Ionescu, 2006). Not only has CSR become essential to improve the image of companies to consumers, it is also helping firms in using their resources more efficiently and in turn improve their productivity as it was stated in the World Economic Forum (WEF, 2015). Nowadays, it is rare to find a company that does not have a CSR annual report, a CSR team, or at least a communication plan dedicated to CSR. For instance, according to the Governance and Accountability Institute, 85% of the S&P 500 companies published CSR Reports in 2017 (Governance & Accountability Institute, 2017).

To manage CSR in supply chains, firms can employ three management strategies

(Ciliberti *et al.*, 2008): (1) Establishing written requirements that suppliers need to respect, these requirements include local laws and CSR international standards, (2) Monitoring the suppliers' performance by conducting surveys and factory inspections and (3) Raising suppliers' awareness through training programs to facilitate their compliance with the company CSR policies. Maintaining CSR in global supply chains is manageable with the help of supplier collaborations (WEF, 2015). Collaborating with suppliers for CSR offers a sustainable business opportunity. It is the constructive partnership between the a company and its suppliers that allows the most significant improvements for the integration of CSR in production processes. Supply Chain CSR policies can help improve a company's revenue by as much as 20% through the resulting economies of scale. For example, by reducing their carbon emission UPS has increased its revenue by 3.4% (UPS CSR Report, 2013) and Nestlé has saved up to 15% (WEF, 2015). Moreover, Unilever has gained 15% in its cost spendings using recyclable alternatives in their products' design (WEF, 2015).

Scholars have also considered CSR from a theoretical perspective. This is noticeable from the special issues that have appeared in the Journal of Management Studies (McWilliams *et al.*, 2006), Academy of Management Review (Bies *et al.*, 2007), and Socio-Economic Review (Brammer *et al.*, 2012). The concept of CSR was accounted for by making use of several concepts ranging from stakeholder theory and institutional theory to the resource-based view and transaction cost economics.

2.2 Text mining and Predictive modeling

With the advances in social networks, data sources and content have grown exponentially. Facing this trend, companies would like to keep as much exploitable information as possible. Nowadays, the majority of data is unstructured and semistructured, such as e-mails, customer calls and news feeds (Kleindorfer *et al.*, 2005). Until a few years ago, this type of data was only stored as it is difficult to explore because of its format. This has changed with the advances in text mining algorithms that made them today's necessity for companies.

Text Mining is concerned with the processing of text corpora to analyze the content and extract knowledge from structured and unstructured text data formats. The main task is the recognition of information and hidden patterns in the document and its interpretation. (Tan *et al.*, 1999) has described text mining as "the process of extracting interesting and non-trivial patterns or knowledge from unstructured text documents". All this is possible thanks to several techniques such as data mining, machine learning, information retrieval, natural language understanding, case-based reasoning, statistics, and knowledge management. The goal is to gain knowledge from large amounts of semi-structured and unstructured texts. In the present work, we have adopted text mining solutions that rely on non-deterministic algorithms to analyze text. It is based on the principles of machine learning, which consists of model training via pre-processed data, called training base, to facilitate future treatments. Many algorithms can be applied to accomplish this task. They can be based on statistics (e.g., Bayesian network), geometry (e.g., k-nearest neighbor), tree of decision (e.g., deep forest), support vector machine and neural networks.

(Tan *et al.*, 1999) describes text mining as a two-stage process. In the first stage known as "text refining", text documents are transformed from a raw-text format into a chosen intermediary format. All along this stage, several actions are performed and they include text extraction and text preprocessing. These tasks are crucial in order to efficiently mine the text, they will be detailed in Chapter 3. In the second stage, referred to as "knowledge distillation", we extract hidden patterns and knowledge from the intermediate form. Intermediate forms can be categorized as either structured (e.g., relational databases) or semi-structured (e.g., conceptual graphs), and document-based or concept-based.

In order to make better decisions and reduce uncertainty, predictive modeling, which is an important tool of text mining can be used (Miner *et al.*, 2012). Predictive modeling is a set of methods for analyzing and interpreting defined data to derive predictions about future events and trends. These predictions give useful information that may assist in decision making. For instance, in our work we have used predictive modeling to predict the CSR risk level for a given supplier by text mining news feeds. This process will be detailed in Chapter 3.

2.3 Supplier Selection and Lot-sizing

In today's competitive environment, companies are looking for efficient ways to optimize their production, maximize their profit and effectively manage their resources. To do so, many purchasing choices that play a significant role in the production and logistics administration need to be taken into account when designing prescriptive supplier selection models.(Aissaoui *et al.*, 2007) list six major purchasing decision processes:(1) make or buy, (2) supplier selection, (3) contract negotiation, (4) design collaboration, (5) procurement, and (6) sourcing analysis. The first step (1) involves taking a decision of whether a company should produce the product/service or outsource it. In the second step, supplier selection, firms select the suppliers from whom they want to procure products and services. This selection stage incorporates multiple criteria and several approaches have been used (Ho *et al.*, 2010) such as the analytic hierarchy process and data envelopment analysis. Step (3) involves the design of the contract between the company and the suppliers and step (4) insures that all requirements are met. In step (6) we consider the evaluation of the total acquisition process.

In this thesis we focus on step (5), procurement, where we answer the question of how to schedule products' orders from suppliers to meet the company's demand while at the same time respecting suppliers capacity and enforcing a CSR policy. Our goal is to minimize the overall procurement cost that includes purchase, ordering and inventory costs. We assume that we are given a list of selected suppliers (from Step 2) and a fixed planning horizon.

Lot-sizing problems are manufacturing scheduling problems. Their main goal is to determine how procurement should take place and gives a solution to how the quantities should be ordered to meet demand while minimizing manufacturing and inventory costs. Many mathematical formulations have been proposed, where each was based on the nature of the addressed problem: having multiple/single suppliers, multiple/single time periods, capacited/incapacitated suppliers. Other problems have added other factors like quality, demand uncertainties...

When developing sustainable supplier selection and lot-sizing models, most researchers have addressed the environmental and economic aspects of sustainability (Azadnia *et al.*, 2015). Incorporating the social criteria in mathematical models has been given more attention in recent years as pressure from governments and social activists have brought more attention to the social aspect in supply chains (Rabieh *et al.*, 2018). However, most of the proposed models were based on hypothetical examples and less attention was given to develop practical models that incorporate supplier selection, lot-sizing and CSR in an integrated mathematical problem (Azadnia *et al.*, 2015). Therefore, to address the above-mentioned problems we propose an integrated sustainable supplier selection and lot-sizing model given CSR risk scores for suppliers. These scores are calculated using machine learning techniques, where we mine newspapers text. In this research work, we are interested in a multi-period inventory lot-sizing scenario, where there are multiple products and multiple suppliers, and products can be carried from one period to another with an additional holding costs. Therefore, we will use (Basnet and Leung, 2005) as a base to build our optimization model, the formulation of our proposed model will be detailed in chapter 4.

2.4 Conclusion

We have reviewed the literature that is closely related to our thesis topic. We have found that CSR has become an important topic and there is an increasing interest in incorporating CSR in decision models. Towards this goal companies are interested in making use of the big data that is increasingly available to them. However, this process may be challenging as models to quantify CSR are not easy to build. Text mining and predictive modeling techniques can help firms process and build knowledge from unstructured data which has become their most used data format. Therefore, these techniques can offer an efficient way to quantify CSR using self-learning algorithms. We have also seen that suppliers selection and lot-sizing optimization models are important because they help companies manage effectively their purchases and maximize their profit. With the availability of big data and predictive modeling power, there is an opportunity to integrate predictive and prescriptive analytics in supplier selection and lot sizing models. This will help companies improve their costs, meet their CSR goals and at the same time gain value out their investment in big data technology.

Chapter 3

A predictive model for detecting supplier non-compliance to corporate social responsibility goals

3.1 Introduction

In order to detect supplier non-compliance in a global supply chain, the use of a validated data is a crucial key factor to effectively monitor suppliers. Having the opinion of experts and collaborating with suppliers to obtain this information is essential, but it is time-consuming and sophisticated strategies need to be developed given the complexity of CSR problems (Widiarto Sutantoputra, 2009). Using secondary data, such as news feed, to collect information about suppliers' behaviour in respecting CSR requirements can solve this issue, as this type of data is easily accessed and quick to find. In this chapter we discuss a predictive model for corporate social responsibility compliance in global supply chains. This model may help prevent another accident like the tragedy of the Rana Plaza factory fire that happened in Bangladesh killing more than 1,110 people and injuring thousands more (White, 2017). Our goal is to text mine news feed to build a model that will serve as an early warning system for supplier's con-compliance with CSR rules; it will predict any potential mishaps at the suppliers end before it becomes the corporate's liability.

In this chapter we describe the predictive model and its different steps. We start with a review of machine learning algorithms, a major component of predictive models. We will then explain how we built our model, from data preprocessing to learning and evaluation.

3.2 Machine Learning

With the increasing abundance of big data, using self-learning algorithms from the machine learning field will help us turn this data into knowledge. Machine learning's goal is to teach machines to carry out tasks by providing them a sample of examples that include instances that have conditions of when a task has to occur or not occur. This involves using algorithms that learn hidden insights from input data to make predictions without the need for explicit programming.

3.2.1 Machine learning phases

A machine learning method includes two phases: learning and prediction. The learning phase consists of learning patterns from data, the intuition behind this is to construct a predictive model. Once the model is built, we move to the prediction step, where we plug new instances into our model to make predictions. The last step is to evaluate the model performance.

3.2.2 Machine learning classes

The major problem classes of machine learning are supervised learning and unsupervised learning.

3.2.2.1 Supervised learning

In the first class, we are given a data set and already know what our correct output should look like, having the idea that there is a relationship between the input and the output. Supervised learning problems are categorized into “classification” and “regression” problems that are used for predicting discrete and continuous values, respectively.

3.2.2.2 Unsupervised learning

In unsupervised learning, we deal with unlabeled data (data that has no sort of meaningful “label” or “class”) and the goal is to extract meaningful information with little or no idea about what our results should look like. Clustering is the most famous unsupervised method, where we use a measurement of similarity between dataset instances. Clustering algorithms use distance-based methods, centroid-based methods, or statistical techniques based on distributions.

In the following sections, we will detail the algorithm class that we have adopted to build our warning system for detecting supplier’s non-compliance.

3.3 Overview of the predictive model

To design a predictive model that can help maintain supplier's compliance with CSR, we need to know what are the reasons that cause suppliers to deviate from CSR guidelines. Recognizing the whys and wherefores of non-compliance can help detect the deviance behaviour of a certain supplier in advance. Based on this knowledge, our predictive model will attribute a CSR risk probability to a supplier. Suppliers who have higher CSR risk probability values are the ones who are likely to be the cause of future damaging incidences in their respective supply chains.

To make our model relevant and practical we attempted to use real life data whenever possible. Given the intricacy of this issue and its impact of public perception, many companies that we have approached for data have refused to cooperate. However, due to pressure from public interest groups, many large firms have adopted some transparency in reporting on their CSR initiatives. We therefore sought to reply on such public reports to validate our model. One such example of firms is Apple, the giant multinational technology company. In terms of suppliers, we focused on Foxconn, a Taiwanese multinational electronics contract manufacturing company that is suppliers the large electronic companies such as Apple, Microsoft, Nintendo, Samsung and Sony. Our choice of Foxconn is also motivated by the fact that they have been found to not obey certain CSR guidelines in past and so their case would be useful for validation to predictability power of our model.

Based on Apple's Supplier Responsibility Progress Report for 2018 (Apple, 2018), CSR issues at the suppliers' end can be categorized into two classes. The first class is labour issues and the second class is health and safety issues. In our model, the labour issues class incorporates issues related to involuntary labour, overtime labour, document falsification, working hours falsification, workers intimidation and discrimination. On the other hand, in the health and safety issues class we address threats that exist in the working environment such as use of illegal chemicals, biological/physical hazards and failure in building structure.

Given the negative impact that supplier non-compliance to CSR can cause to a company's financial well being as well as public perception (Sills Egsgard LLP Article, nd), having an early warning system would be a valuable tool. Companies can setup formal supplier relationship management procedures to collect CSR data, incorporate it with other data, and analyze it for possible supplier deviance. However, for different reasons, such as supplier not reporting accurate data or complex supply network where lower tier suppliers are not known to the company, supplier noncompliance can still happen and sometimes with deadly and costly consequences (e.g., the death of more than 1,110 people and the injury of thousands more in the Rana Plaza building collapse and the negative publicity for its clothing retailers (Anisul Huq *et al.*, 2014)). To overcome this challenge, in this thesis we propose to use secondary data, such as news feed, to gather intelligence on local suppliers' CSR performance.

Our goal is to text mine news feed and classify suppliers according to predefined CSR categories and use them to identify risky suppliers. We will classify news items into one of the three classes: (1) news items that report on labour CSR issues, (2) news items that report on health and safety CSR issues, and (3) news items that do not report on any CSR issues. As a result, the final output of the predictive model will be three values (x, y, z), where:

- x: probability news item belongs to labour issues class

- y: probability news item belongs to health and safety issues class
- z: probability news item belongs to no issues class

We will use Factiva, global news database, our source for data. We will employ text mining techniques and classification algorithms to classify each news item into one of the three classes.

3.4 Predictive model components

The main steps of building a predictive model are data collection and transformation, data preprocessing and model learning and evaluation. A flowchart illustrating these major steps, as applied to our problem, is shown in Figure 3.1.

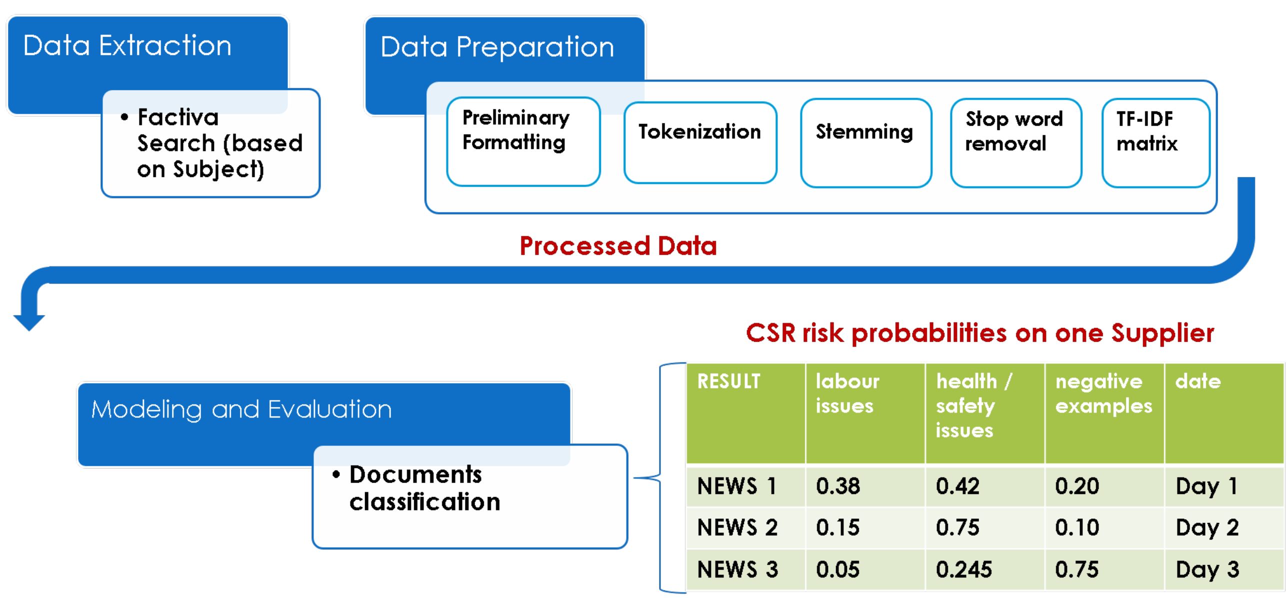


Figure 3.1: Flowchart of the main steps of the predictive model.

First, we extract data and transform it. During this stage, we retrieve newspapers from Factiva and then run some data preparation methods. In the preprocessing stage, we use text mining methods such as data cleaning and term extraction. The resulting data can now fit into the self-learning algorithm and go to the next stage which is the learning and evaluation phase. In this step the machine learning algorithm will search for hidden insights and perform the classification. This learning algorithm is chosen and evaluated based on the desired output and the quality of the prediction. The final output will be probabilities of the news item belonging to each of the three classes introduced previously.

3.4.0.1 Data Extraction

Our data was collected from Factiva, a database of newspaper articles from around the world, where we searched for news related to a specific supplier. Using Factiva's search tool, we searched the news for three different subjects: labour issues, health and safety issues and no issues subject. We restricted our search to news articles related to Foxconn company and the region to China, since most of Foxconn's factories are in that region.

In Table 3.1, we give a summary of our case study of Apple and Foxconn, where we mention for each class how many news articles we have used, the subjects used in Factiva search tool and the time periods [1].

3.4.1 Data Preparation

Data preparation consists of tow steps: data transformation and data preprocessing. Theses steps are detailed in the following sections.

[1] At the day search for news articles was conducted, we put all dates for labour issues class and health and safety class. For negative examples, we put three months and we have chosen the first 100 articles. In the Time period column we present the resulting date ranges.

Classes	Time periods	Articles count	Language	Country	Subject
Labour issues class	27 July 2009 - 22 March 2018	115	English	China	Labour Disputes Or Workers pay Or Workplace Discrimination/Abuse Not Workplace Safety/Health Issues
Health and safety issues class	22 July 2009 - 22 August 2016	43	English	China	Workplace Safety/Health Issues Not Workers pay Not Workplace Discrimination/Abuse Not Labour Disputes
Negative examples class	2 March 2017 - 23 May 2018	100	English	China	Not any subject related to Labour issues class and Health and safety issues class

Table 3.1: Summary table of Apple and Foxconn news articles.

3.4.1.1 Data Transformation

To build analytical models and extract consistent knowledge, the data needs to have a specific format other than raw text. The transformation step allows as to convert the raw data into a form that is more amenable to text mining algorithms. To do so, we have developed a script that parses the plain text and transforms it into a structured data. The final file will be a document matrix, where each row is a document and its properties, such as tilte, main text, date and the class it belongs to, are the columns. In the 'class' column we classify news items into one of the three classes: Class 0 has news items that report on labour CSR issues, class 1 for news items that report on health and safety CSR issues, and class -1 for news items that do not report on any CSR issues. After performing the preliminary formatting, the data output's format is illustrated in Table 3.2.

	Title	Text	Date	Class
NEWS 1	Title1	Text1	Date 1	0
...	...	...	...	...
NEWS n	Title n	Text n	Date n	-1

Table 3.2: Data Transformation.

3.4.1.2 Data Preprocessing

Text analysis is performed on terms that construct the news document, therefore, the text needs to be preprocessed to obtain adequate terms. To do so, we first perform tokenization that consists of chopping the document into tokens, or words. Then, we perform the following tasks: remove numbers, erase punctuation, convert all text to lower case, remove stop words: 'at', 'the', 'a', 'in', as they do not provide any significant information. After this step, we do stemming, which involves converting each word to its base word, and the resulting terms are extracted, and text analysis is performed using their term frequency and inverse document frequency (TF-IDF) matrix. With this step we seek to grant lexical relevance of a term within a document. Using the TF-IDF matrix, we apply a relationship between a document and a set of other documents sharing similarities in keywords. In a way, we are looking for a relationship of quantity and lexical quality across a set of documents. For a query with a term x, a document is more likely to be relevant as a response to the query, if that document has a certain occurrence of that term within it, and that term has a rarity in other documents connected to the first. The mathematical formula for calculating the TF-IDF of a term x in a document y is:

$$W_{x,y} = tf_{x,y} * \log(\frac{N}{df_t}) \tag{3.1}$$

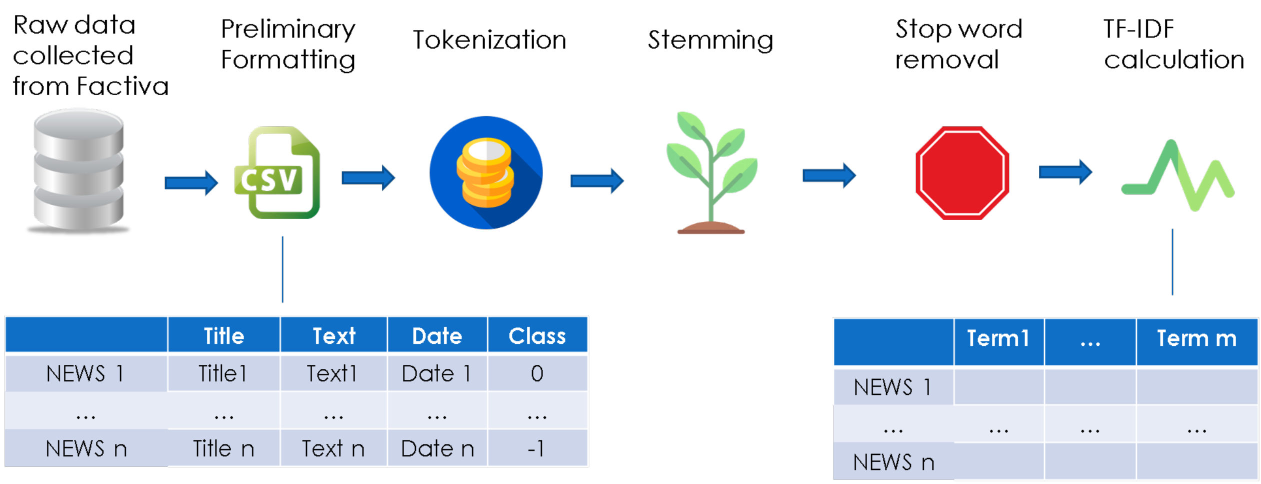
where:

$tf_{x,y}$**:** is number of occurrences of the term x within document y

df_t**:** is number of documents containing term x

N: is total number of documents

To summarize, the final process of data preparation is presented in Figure 3.2.



	Title	Text	Date	Class
NEWS 1	Title1	Text1	Date 1	0
...	...	...	...	...
NEWS n	Title n	Text n	Date n	-1

	Term1	...	Term m
NEWS 1			
...	...	...	...
NEWS n			

Figure 3.2: Data Preparation.

3.4.2 Predictive models

In this section, we present the self-learning algorithms we used to train our predictive model. The choice of these algorithms was motivated by our desired output. Since we want each news article to have a probability to belong to each of the three classes, we have used algorithms that return the probability of the samples for each class in the model. For algorithms implementation, we have used the open source library scikit-learn (Buitinck *et al.*, 2013).

We start by Random Forest (RF) which is an algorithm that was formally proposed by (Breiman, 2001). This algorithm combines random subspace concepts and bagging. RF performs based on multiple decision trees trained on subsets of slightly different data and uses adaptive strategies (boosting) or random (bagging). The idea is to aggregate a large number of models while avoiding over-fitting. The RF algorithms form a family of classification methods that relies on the combination of several decision trees. Each decision tree is composed from a random combination of explanatory variables (features), the final decision is the result of previously built voting decision trees. The second algorithm is Linear Support Vector Classifier (SVC). This algorithm is similar to Support Vector Machines (SVMs), but having a linear kernel. They are effective when the number of features is high and when the dimension is greater than the number of samples. Therefore, this algorithm is a good option since the TF-IDF matrix has a high number of news article terms (features) compared to the number of samples. SVMs construct a hyperplane or a group of hyperplanes to perform several tasks like classification, regression and outliers detection. Having an infinity of hyperplanes separators, the optimal hyperplane is defined as the hyperplane that maximizes the margin between the samples and the separating hyperplane. The third algorithm is the Multinomial naive Bayes classifier (MultinomialNB). It is an instance of the naive Bayes algorithm and it is based on the independence assumption of all features. MultinomialNB is used for multinomially distributed data and its performance in classification is competitive (Zhang, 2004a) especially for discrete features classification. Since text-mining is based on fractional word count TF-IDF, this algorithm is a good option for building our predictive model. Lastly, Stochastic Gradient Descent (SGD) classifier is an estimator that implements regularized linear models with SGD learning. This algorithm is efficient, easy to implement and intensively applied for large scale data (Zhang, 2004b). Moreover it is highly applied in building text classification and natural language processing predictive models [2].

[2] The official website of scikit-learn offers interesting and practical examples to develop the stated algorithms: `https://scikit-learn.org/stable/index.html`.

3.4.3 Model evaluation

One of the common performance measures for predictive models is the confusion matrix (see Figure 3.3 and Table 3.3).

Each matrix column represents the number of occurrences of an estimated class, while each row represents the actual observed number. Unlike simple aggregate error rates, the confusion matrix summarizes the correct and incorrect predictions for each class. The diagonal entries show the correct predictions for each class. The confusion matrix can also provide performance measures such as precision, recall (or sensitivity) and f1-score.

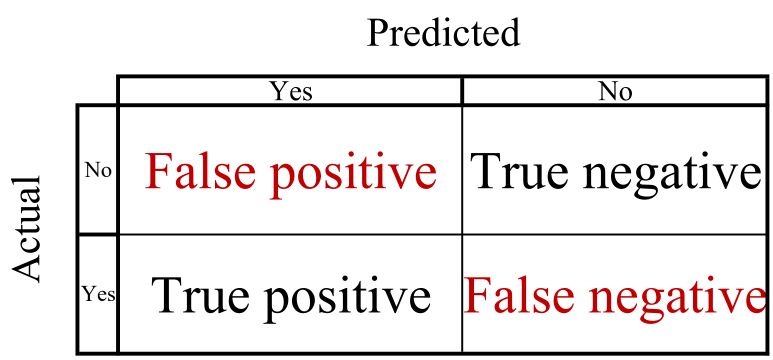


Figure 3.3: Confusion Matrix.

Case Name	Description
True positive (TP)	Correctly predicted class “Yes” events
False positive (FP)	Incorrectly predicted class “Yes” events
True negative (TN)	Correctly predicted class “No” events
False negative (FN)	Incorrectly predicted class “No” events

Table 3.3: Explanation of TP, FP, TN and FN.

Precision is defined as the number of true positives (TP) divided by the sum of true positives and false positives (FP), i.e., the ratio of correctly predicted positive observations to the total predicted positive observations. $Precision = \frac{\text{TP}}{\text{TP + FP}}$.

In our case, precision of class i = $\frac{\text{number of documents correctly attributed to class i}}{\text{number of documents attributed to class i}}$.

Recall is defined as the number of true positives divided by the sum of the true positives and false negatives, i.e., the ratio of correctly predicted positive observations to the total of correctly predicted events. $Recall = \frac{\text{TP}}{\text{TP + FN}}$.

In our case recall of class i = $\frac{\text{number of documents correctly attributed to class i}}{\text{number of documents belonging to class i}}$.

The f1-score is the weighted average of precision and recall, f1-score reaches its best value at 1 (perfect precision and recall) and worst at 0. F1-score of class i $= \frac{\text{2*(Recall * Precision)}}{\text{Recall + Precision}}$.

Once we calculate the TF-IDF values, we input this matrix into the learning algorithm to perform the classification. In order to pick the right algorithm for our modeling phase, we have tested several classification algorithms on our dataset. The chosen algorithm is the one that offers the best accuracy. A summary of the tested algorithms' accuracy is presented in Figure 3.4.

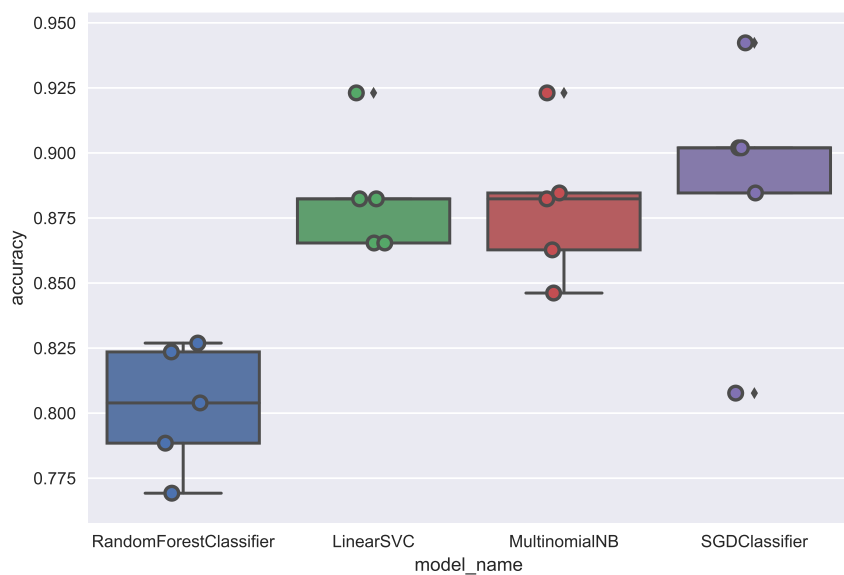


Figure 3.4: Data Evaluation.

As we can see the SGD classifier gave the highest accuracy. To evaluate the performance of the classifier, we have generated the confusion matrix for our three classes: health and safety issues class, labor issues class and negative examples class (documents containing subjects other than health and safety issues and labor issues). The confusion matrix is shown in Figure 3.5.

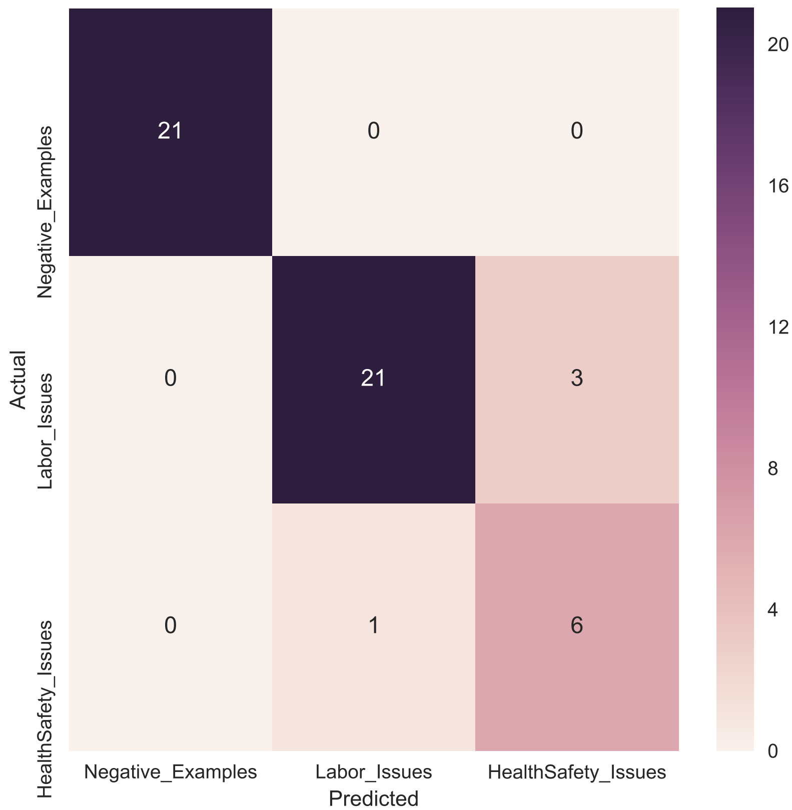


Figure 3.5: Confusion matrix of SGD classifier.

In Table 3.4, we report on different performance measures for the model. In that table support is the number of news articles used for evaluating the corresponding class.

Classes	precision	recall	f1-score	support
labour issues class	0.95	0.88	0.91	24
health and safety issues class	0.67	0.86	0.75	7
negative examples class	1.00	1.00	1.00	21
avg / total	0.93	0.92	0.93	52

Table 3.4: Confusion matrix of SGD classifier.

We see that the SGD classifier has correctly classified all instances from *negative examples*, but had some mis-classifications for *labor*, and *health and safety issues* classes, which is expected as these subjects have many terms in common. The mis-classification does not present an issue as the the final risk probability for a supplier will be the sum of the probabilities of belonging to *labor issues* and *health and safety issues* classes.

In Table 3.5 we illustrate our final result on how the CSR risk probability is calculated. We denote by (a_k, b_k, c_k) the probabilities for a document x_k to belong to each of the

	Labor issues	*Health and Safety issues*	*Negative Example*	CSR risk probability per document	CSR risk probability of a supplier
$newsx_1$	a_1	b_1	c_1	$r_1 = a_1 + b_1$	$Pr = \frac{\sum_{k=1}^{N} r_k}{N}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
$newsx_N$	a_N	b_N	c_N	$r_N = a_N + b_N$	

Table 3.5: Calculating CSR risk probability of a supplier.

three classes, where k is an integer in $[1, N]$ and N is the total number of news documents. First we calculate the CSR risk probability , r_k, for each document. This value is the sum of the probabilities that news document $newsx_k$ belongs to the *Labor issues* and *Health and Safety issues* classes. The total CSR risk probability of a supplier will be the average of CSR risk probabilities of each document and it is equal to $Pr = \frac{\sum_{k=1}^{N} r_k}{N}$. A summary for our Foxconn and Apple case is in Appendix A.1. The resulting CSR risk probability for Foxcoon is $Pr = 0.56$

3.5 Conclusion

We have presented the different steps to build the predictive model for corporate social responsibility using machine learning. After explaining the process of data collection and preprocessing we introduced the machine learning classification algorithms. We evaluated each algorithm based on its accuracy and SGD classifier has given the best results. The output of the algorithm was used to calculate the CSR risk probability of a supplier. Theses probabilities will be used in the following chapter, where we will present a prescriptive model to optimize supplier selection and order lot sizing subject to CSR requirements.

Chapter 4

A prescriptive analytics models for supplier selection and lot sizing that incorporates CSR

4.1 Introduction

In the present chapter, results from the previous predictive model are used to create a supplier selection optimization model that ensures CSR compliance in global supply chains. We will first present the formulation of this optimization model and comment on its structural properties. Then we propose a method for generating feasible problem instances and another for solving the optimization problem. Finally, we present our computational results and managerial insights.

4.2 Problem formulation

Basnet and Leung (2005) have presented in their extension a multi-period lot-sizing and supplier selection problem. Here we propose to extend this model by incorporating supplier production capacity and CSR requirements. Using the following notation, we introduce our capacitated multi-period lot-sizing and supplier selection problem under a corporate social responsibility constraint.

Indices:

$i = 1 \ldots I$ index of products.

$j = 1 \ldots J$ index of suppliers.

$t = 1 \ldots T$ index of time periods.

Parameters:

D_{it} = demand of product i in period t.

P_{ij} = purchase price of product i from supplier j.

H_i = holding cost of product i per period.

O_j = ordering cost from supplier j.

C_{ji} = maximum number of units of product i that can be ordered from supplier j per period.

Pr_j = CSR risk probability for supplier j.

α a threshold for CSR constraint, $0 < \underline{\alpha} \leq \alpha \leq \overline{\alpha} \leq 1$, $\underline{\alpha}$ and $\overline{\alpha}$ are the minimum and maximum tolerance for CSR non-compliance, respectively.

$\gamma \geq 1$ capacity flexibility parameter.

ω pricing parameter, we assume $0.1 \leq \omega \leq 0.2$.

β holding cost parameter, we assume $0.1 \leq \beta \leq 0.2$.

Decision variables:

X_{ijt} = quantity of product i ordered from supplier j in period t.

$Y_{jt} = 1$ if an order is placed to supplier i in time period t, 0 otherwise.

The problem can be formulated as follows:

$$\min \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=1}^{T} P_{ij}X_{ijt} + \sum_{j=1}^{J}\sum_{t=1}^{T} O_j Y_{jt} + \sum_{i=1}^{I}\sum_{t=1}^{T} H_i(\sum_{k=1}^{t}\sum_{j=1}^{J} X_{ijt} - \sum_{k=1}^{t} D_{ik}) \tag{4.1}$$

$$\text{s.t} \quad \sum_{k=1}^{t}\sum_{j=1}^{J} X_{ijt} - \sum_{k=1}^{t} D_{ik} \geq 0, \quad \text{for all } i \text{ and } t, \tag{4.2}$$

$$C_{ji}Y_{jt} - X_{ijt} \geq 0, \quad \text{for all } i,\ j \text{ and } t, \tag{4.3}$$

$$\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=1}^{T} Pr_j X_{ijt} Y_{jt}}{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=1}^{T} X_{ijt} Y_{jt}} \leq \alpha, \tag{4.4}$$

$$Y_{jt} = \{0, 1\}, \quad \text{for all } j \text{ and } t, \tag{4.5}$$

$$X_{ijt} \geq 0, \quad \text{integer for all } i,\ j \text{ and } t. \tag{4.6}$$

The objective function in (4.1) represents the cost we want to minimize, it includes three costs: 1) cost of the purchased products, 2) cost of ordering from suppliers, and 3) inventory holding cost. Constraint set (4.2) enforces that demand is met in each period. In other words, inventory is always nonnegative so that we are not allowing for backlogs. Constraint (4.3) ensures that we do not place an order that exceeds a supplier's capacity. Constraint set (4.4) ensures that that the average risk for all products orders from all suppliers in all periods should not exceed a threshold α. The value of α varies from 0 to 1, when it is closer to 0, it implies that the company would like to enforce more strict CSR standards. Constraint set (4.5) and (4.6) enforce the binary and nonnegativety requirements for variables X_{ijt} and Y_{jt}, respectively.

Because of the nonlinearity of constraint (4.4), problem ((4.1)–(4.6)) is a mixed integer nonlinear program, which are generally very hard to solve. Luckily it can be converted to a mixed integer program. To do so we first note that constraint (4.4) is equivalent to $\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=1}^{T} Pr_j X_{ijt} Y_{jt} \leq \alpha \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=1}^{T} X_{ijt} Y_{jt}$. Although we got rid of the fractional terms, we still have the bilinear terms $X_{ijt}Y_{jt}$. To linearize these bilinear terms we note that when Y_{jt} is 0 then X_{ijt} should be 0, given the capacity constraint (4.3). On the other hand when $Y_{jt} = 1$ then $X_{ijt}Y_{jt} = X_{ijt}$. Therefore, we can simply replace the product $X_{ijt}Y_{jt}$ with X_{ijt}.

Reformulating ((4.1)-(4.6)) we obtain the following mixed integer program (MIP):

min(4.1)

s.t. (4.2), (4.3), (4.5), (4.6)

$$\sum_{i}\sum_{j}\sum_{t}(Pr_j - \alpha)X_{ijt} \leq 0 \tag{4.7}$$

4.3 Problem Analysis

The lot sizing problem goes back to the famous work by Wagner and Whitin (1958) where they considered a single product and multi-period problem. This was extended in many different ways. The extensions that we are interested in are those that included capacity (e.g., Florian and Klein 1971) and multiple products and suppliers (e.g., Basnet and Leung 2005). When studying these problems it is customary in the literature to investigate their structural properties. In this regard, three main properties have been investigated: the *zero-inventory policy* , the *single-sourcing policy* and the *capacity constrained order sequence*. Below we define each of these policies and show that, unfortunately, they do not hold for our problem, indicating the complexity of our problem.

Zero Inventory property

Definition 1 (Zero Inventory Policy)**.** *In a zero inventory policy we have* $(\sum_{k=1}^{t}\sum_{j=1}^{J} X_{ijt} - \sum_{k=1}^{t} D_{ik})X_{ijt} = 0$ *for all i, t, and j; i.e., there is no period where an order is placed while at the same time we have a positive inventory in that period.*

As in Definition 1, the zero inventory policy implies that we may place an order for a product only if its inventory is zero. We will use a counter example to show that this property does not hold for our problem. Consider the instance of our problem shown in Table 4.1. The solution for product 5 is shown in Table 4.2. We see clearly that orders were placed even though there is inventory from previous time periods to meet cumulative demand.

	Capacity	Products				
	C_{ji}	i=1	2	3	4	5
Suppliers	j=1	4	11	3	6	7
	2	6	6	12	4	8
	3	8	5	8	5	5
	4	7	1	6	6	7
	5	7	2	7	8	4

	Price	Suppliers				
	P_{ij}	j=1	2	3	4	5
Products	i=1	45	26	45	65	33
	2	35	21	53	5	44
	3	57	55	29	23	57
	4	30	62	49	33	18
	5	38	16	73	27	6

	Demand	Time periods				
	D_{it}	t=1	2	3	4	5
Products	i=1	18	25	14	15	15
	2	13	20	10	15	13
	3	24	12	18	17	10
	4	16	21	15	25	17
	5	11	23	24	25	11

Ordering Price	Suppliers				
	j=1	2	3	4	5
O_j	53	189	60	146	122

Holding Price	Products				
	i=1	2	3	4	5
H_i	7.02	5.18	7.25	6.30	5.25

CSR risk probability	Suppliers				
	j=1	2	3	4	5
Pr_j	0.68	0.58	0.26	0.60	0.12

Table 4.1: Data for counter example for zero inventory property ($I = J = T = 5$).

Optimal cost = 14552.63

X_{ijt}		Time periods				
Products	Suppliers	t=1	2	3	4	5
i=1	j=1	4	4	1	2	2
	2	6	6	6	6	6
	3	1	8	0	0	0
	4	0	0	0	0	0
	5	7	7	7	7	7
	Demand	18	25	14	15	15
	Order	18	25	14	15	15
	Inventory	0	0	0	0	0
i=2	j=1	8	11	3	8	6
	2	6	6	6	6	6
	3					
	4	1	1	1	1	1
	5					
	Demand	13	20	10	15	13
	Order	15	18	10	15	13
	Inventory	2	0	0	0	0
i=3	j=1					
	2	10		2	3	
	3	8	8	8	8	4
	4	6	6	6	6	6
	5					
	Demand	24	12	18	17	10
	Order	24	14	16	17	10
	Inventory	0	2	0	0	0
i=4	j=1	6	6	6	6	6
	2					
	3					
	4	3	6	6	6	3
	5	8	8	8	8	8
	Demand	16	21	15	25	17
	Order	17	20	20	20	17
	Inventory	1	0	5	0	0
i=5	j=1			1	6	
	2	8	8	8	8	7
	3					
	4	7	7	7	7	
	5	4	4	4	4	4
	Demand	11	23	24	25	11
	Order	19	19	20	25	11
	Inventory	8	4	0	0	0

Table 4.2: Solution to counter example for zero inventory property ($I = J = T = 5, \alpha = 0.5$).

Single-sourcing policy

Definition 2 (Single sourcing policy)**.** *In a single sourcing policy we have* $X_{ijt}X_{ikt} = 0$ *for all* $i, t, j \neq k$*, i.e., we cannot order from more than one supplier in the same period.*

As in Definition 2, the single-sourcing policy implies that we may place an order for a product only from one supplier in the same period. We will use a counter example to show that this property does not hold for our problem. Consider the same instance of our problem shown in Table 4.1. The solution for product 5 is shown in Table 4.2. We see clearly that, in the same time period orders are placed from multiple suppliers.

Capacity constrained order sequence

Definition 3 (Capacity constrained order sequence, Florian and Klein (1971))**.** *An order sequence is capacity constrained if for a given product* i *and supplier* j, $SC = \{t', \;\; such \;\; that \;\; 0 < X_{ijt'} < C_j\}$, $|SC| <= 1$ *and* $X_{ijt} = 0$ *or* $X_{ijt} = C_j$ *for all* t $(\neq t' \;\; when \;\; |SC| = 1)$*, i.e., there is at most one period for which we order a positive amount that is not equal to a supplier's capacity and in all other period we either order nothing or order at supplier's capacity.*

To show that this property does not hold for our problem consider the problem instance in Table 4.3. The solution to that instance is shown in Table 4.4. We see that for product 3 we order from supplier 4 more than once at an amount that is below that supplier's capacity.

Capacity		Products				
	C_{ji}	i=1	2	3	4	5
Suppliers	j=1	15	18	10	29	18
	2	22	15	26	22	18
	3	20	18	15	9	11
	4	12	25	7	21	30
	5	14	7	25	26	23

Price		Suppliers				
	P_{ij}	j=1	2	3	4	5
Products	i=1	45	26	45	65	33
	2	35	21	53	5	44
	3	57	55	29	23	57
	4	30	62	49	33	18
	5	38	16	73	27	6

Demand		Time periods				
	D_{it}	t=1	2	3	4	5
Products	i=1	18	25	14	15	15
	2	13	20	10	15	13
	3	24	12	18	17	10
	4	16	21	15	25	17
	5	11	23	24	25	11

Ordering Price	Suppliers				
	j=1	2	3	4	5
O_j	53	189	60	146	122

Holding Price	Products				
	i=1	2	3	4	5
H_i	7.02	5.18	7.25	6.30	5.25

CSR risk probability	Suppliers				
	j=1	2	3	4	5
Pr_j	0.68	0.26	0.12	0.60	0.58

Table 4.3: Problem instance data for counter example for capacity constrained production sequence policy ($I = J = T = 5$)

Final cost = 9430.24

X_{ijt}		Time periods				
Products	Suppliers	t=1	2	3	4	5
i=1	j=1					
	2		22			
	3	4			2	
	4					
	5	14	3	14	14	14
i=2	j=1					
	2					
	3					
	4	18	25		15	13
	5					
i=3	j=1					
	2					
	3	15	8	15	13	
	4	7	7		7	7
	5	2				
i=4	j=1					
	2					
	3					
	4					
	5	16	21	15	25	17
i=5	j=1					
	2					
	3					
	4					
	5	14	23	23	23	11

Table 4.4: Solution to counter example for capacity constrained production sequence policy ($I = J = T = 5, \alpha = 0.8$)

Finally we show in Theorem 1 that problem (4.1)-(4.7) is $\mathcal{NP}$_Hard.

Theorem 1. *Problem (4.1)-(4.7) is $\mathcal{NP}$_Hard.*

Proof. We will show that problem (4.1)-(4.7) can be reduced to the knapsack problem which is known to be $\mathcal{NP}$_Hard (Toker and Ozbay, 1995). To do so consider the special case where $I = 1, T = 1, O_j = H_1 = 0$ for all i, and $C_j = \infty$ for all $j = 1, \ldots, J$, i.e., we consider an uncapacitated single product and period supplier selection and lot sizing problem. Thus problem (4.1)-(4.7) reduces to

min $\sum_{j=1}^{J} P_{1j} X_{1j1}$

s.t. $\sum_{j=1}^{J} (Pr_j - \alpha) X_{1j1} \leq 0$

$X_{1j1} \geq 0$ and integer

which is a knapsack problem. □

4.4 Heuristic

As we have shown in Theorem 1, problem (4.1)-(4.7) is $\mathcal{NP}$_Hard and therefore it cannot be solved efficiently unless $\mathcal{P}$=$\mathcal{NP}$. Our goal in this section is to develop an effective heuristic method to solve problem (4.1)-(4.7). The idea is to decompose the problem and solve the subproblems efficiently. The heuristic is described in Algorithm 1.

Algorithm 1 Heuristic to solve optimization problem

for each product **do**
    Sort suppliers by product prices
    Find minimum number of suppliers to meet demand in all time periods
    if demand exceeds capacity in a time period t **then**
        Execute lot-shifting technique to get a feasible demand solution
    else
        Start by supplier having the lowest price, exhaust its capacity and move to the next until demand is met in all time periods
    end if
end for
Compute $sum_{Pr} = X(i,j,t) * Y(j,t) * (Pr(j) - \alpha)$ for all products
if $sum_{Pr} \leq 0$ **then**
    Finish
else
    repeat(Choose product with the most positive sum_{Pr})
        Sort suppliers by CSR risk probabilities
        Find minimum number of suppliers to meet demand in all time periods
        if demand exceeds capacity in a time period t **then**
            Execute lot-shifting technique to get a feasible demand solution
        else
            Start by supplier having the lowest CSR risk probability, exhaust its capacity and move to the next until demand is met in all time periods
        end if
    until $sum_{Pr} \leq 0$
end if

4.4.1 Heuristic detailed procedures

To execute the heuristic procedure, we have developed three functions (see Algorithm 2):

1. **Min_Supplier_Number**, this function returns the minimum number of suppliers needed to meet demand in all time periods.

2. **Feasible_Demand**,this function tests if we have a feasible solution to start

with. If not, we would perform a lot shifting technique to fix the infeasibilty. The lot shifting method back-shifts the excess demand to some prior period so demand will not exceed capacity in any time period.

3. **Solve_Model**, this function will solve the optimization model based on a given criteria. The parameter criteria can take two values: price or risk. If price was passed as a parameter the function **Solve_Model** will sort suppliers by their products' prices. For each product, it will start by ordering from the supplier with the lowest price, exhausts its capacity and moves to the next until demand is met in all periods. If risk was passed as a parameter, the function will instead sort the suppliers by their CSR risk probabilities and start placing orders from suppliers having the lowest CSR risk probabilities.

Algorithm 2 Algorithm to solve model with CSR constraint

function MIN_SUPPLIER_NUMBER(product,C,D)
    $CumSumC$:=Cumulative sum of capacity C of product for all t
    $CumSumD$:=Cumulative sum of demand D of product for all t
    for $j = 1$ to $size(cumSumC)$ **do**
        $k := 0$
        for $t = 1$ to $size(cumSumD)$ **do**
            if $CumSumD_t \leq cumSumC_j{*}t$ **then**
                $k := k + 1$
                if $k = size(cumSumD)$ **then**
                    return $j + 1, cumSumC_j$
                end if
            end if
        end for
    end for
end function
function FEASIBLE_DEMAND($D_i, total\_Capacity$)
    $temp\_D = D_i$
    for $t = size(D_i)$ to 1 **do**
        if $temp\_D_{it} \geq total\_Capacity$ **then**
            $aux := temp\_D_{it}$ - $total\_Capacity$
            $temp\_D_{it}$:=total_Capacity
            $temp\_D_{it-1}$:=$temp\_D_{it-1}$+aux
        end if
    end for
    return $temp\_D$
end function

```
function SOLVE_MODEL(criteria)
    for i = 1 to I do
        Sort suppliers in Pr, P, O, C based on criteria
        nbSupp, total_Capacity :=Min_Supplier_Number(i,C,D)
        Take first nbSupp rows from P, C and O
        temp_C = C
        for t = 1 to T do
            product_cost_per_period = 0
            ordering_cost_per_period = 0
            D_perPeriod = Feasible_Demand(D_i, total_Capacity)
            j := 1
            while D_perPeriod ≥ 0 do
                D_perPeriod = D_perPeriod − temp_C_{i,j}
                if D_perPeriod ≤ 0 then
                    temp_C_{i,j} = D_perPeriod + temp_C_{i,j}
                end if
                X(i, j, t) := temp_C_{i,j}
                product_cost_per_period = product_cost_per_period+temp_C_{i,j} ∗ P_{i,j}
                ordering_cost_per_period = ordering_cost_per_period + O_j
            end while
            return X(i, j, t), Y(j, t)
        end for
    end for
end function
X(i, j, t), Y(j, t) = Solve_Model(Price)
Compute sum_Pr = X(i, j, t) ∗ Y(j, t) ∗ (Pr(j) − α) for all products
if sum_Pr ≤ 0 then
    Finish
else
    repeat (Choose product with the most positive sum_Pr)
        Solve_Model(Risk)
    until sum_Pr ≤ 0
end if
```

We have tested our Heuristic on 216 instances that were generated as describe in the previous sections. We have tested it on problems having $\alpha = 0.3$ and $\gamma = 1$. From the 216 problems, our Heuristic could not find solutions that respect the CSR

constraint for 4 instances. The infeasibility arises from the CSR constraint. The following example illustrates the issue.
Having a capacity of 23, the following vector gives a feasible ordering scenario vector that respects capacity for all time periods but not the CSR constraint.

$$[\ 23\ ,\ 13\ ,\ 21\ ,\ 21\ ,\ 13\ ,\ 23\ ,\ 23\ ,\ 12\ ,\ 23\ ,\ 23\ ,\ 21\ ,\ 23\ ,\ 14\ ].$$

In this vector we see that the maximum capacity is reached several times, which means that we ended up ordering from suppliers having high CSR risk probabilities multiple times, which gives a solution that does not respect the CSR constraint. To overcome this issue, we need to spread ordering over time periods so maximum capacity would not be reached as often as in the previous solution, this way we are will carry more products inventory between time periods but we will avoid ordering from suppliers having high CSR risk probabilities. In other words we need to minimize the difference between the maximum and minimum values in the ordering vector.
The new ordering vector that respects both capacity and CSR constraints will be:

$$[\ 23\ ,\ 18\ ,\ 18\ ,\ 19\ ,\ 18\ ,\ 18\ ,\ 18\ ,\ 18\ ,\ 18\ ,\ 18\ ,\ 20\ ,\ 23\ ,\ 14\ ].$$

We see that demand is met in all time periods, capacity is respected and we did not use the maximum capacity as often as in the previous solution, which means that we ordered less from suppliers having high CSR risk probabilities. Our proposed solution to minimize the difference between the maximum and minimum values will use a value k which is given to reduce this difference. For instance, given a scalar k, we will always prefer ordering a quantity k for a time period and only order more than k to meet the demand for time period. In the previous example, we see that in most time periods the ordered quantity is 18 and the cases where we ordered more were only to meet the period demand. This procedure will ensure reducing the quantities ordered from suppliers having a high risk and as a result respect the CSR risk constraint. In our computations we have chosen the value k in the range :

$$[\, med - 2 \, , med - 1 \, , med \, , med + 1 \, , med + 2 \,]$$

Where med is the median value of all unique numbers in the ordering vector. For each product, we will minimize the difference between the ordering vector values using each time a value of k from the range stated above. After performing this procedure, we will have 4 ordering scenario vectors, each, for a given value of k. In Algorithm 3 we describe the steps for choosing the ordering vector:

Algorithm 3 Heuristic extension to solve CSR infeasible problems

for a given product **do**:
    temp_Demand1 = MinDiff1(Demand) for k_1
    temp_Demand2 = MinDiff2(Demand) for k_2
    temp_Demand3 = MinDiff3(Demand) for k_3
    temp_Demand4 = MinDiff4(Demand) for k_4

    Ordering the suppliers by the lowest risk and ordering first from suppliers having lowest risk, exhaust their capacity and move to the next, we will end up with 4 ordering scenarios:

    X_inter1, X_inter2, X_inter3, X_inter4

    We will calculate the CSR risk constraint for each of the 4 solutions for the present product

    $sum_{Pr1} = X\_inter1(j,t) * Y(j,t) * (Pr(j) - \alpha)$
    $sum_{Pr2} = X\_inter2(j,t) * Y(j,t) * (Pr(j) - \alpha)$
    $sum_{Pr3} = X\_inter3(j,t) * Y(j,t) * (Pr(j) - \alpha)$
    $sum_{Pr4} = X\_inter4(j,t) * Y(j,t) * (Pr(j) - \alpha)$

    The chosen ordering scenario will be the one that gives $min = (\, sum_{Pr1} \, , sum_{Pr2} \, , sum_{Pr3} \, , sum_{Pr4} \,)$
end for

4.4.2 Heuristic results and analysis

The heuristic results and how they compare with CPLEX are presented in Appendix A.4. We have reported for each instance the number of products, number of suppliers, number of time periods, cost produced by CPLEX, cost produced by our Heuristic, time spent by the Heuristic to solve the problem and a cost ratio. This ratio is calculated to compare the cost values between the Heuristic and CPLEX and it is computed as follows:

$$ratio = \frac{cost_{Heuristic} - cost_{CPLEX}}{cost_{CPLEX}}$$

In Figure 4.1 we show both the heuristics and CPLEX cost data for the all 216 problem instances.

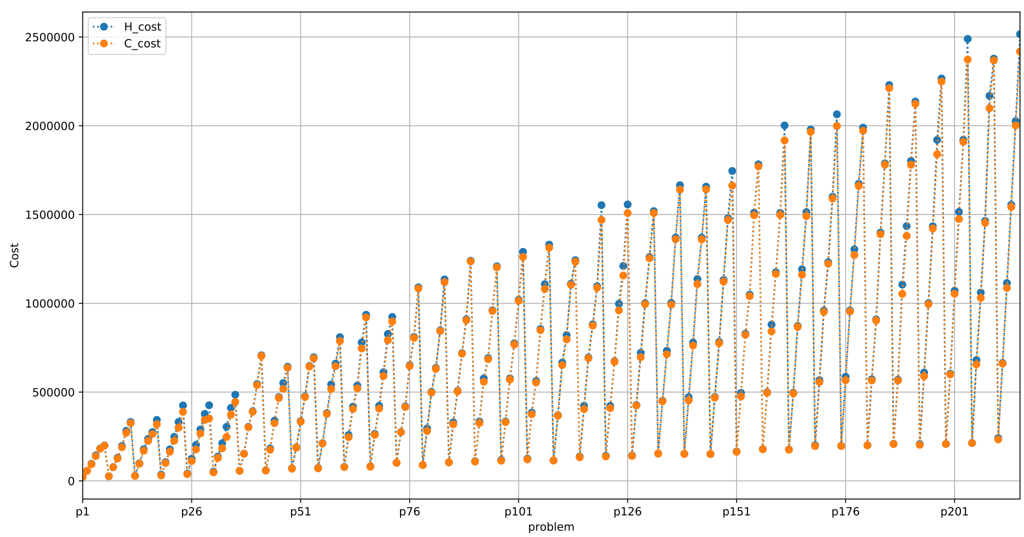


Figure 4.1: Comparison of cost solutions from the Heuristic and CPLEX

As we can see the heuristic is giving approximately the same solutions as CPLEX. On average the heuristic performance ration is 2.74%, i.e., the heuristic objective value are on average within 2.74% of the CPLEX optimal values.

Overall, the heuristic is performing well, however it remains computationally expensive as we can see in the time spent to find the solution. However, we note that our heuristic shows comparable time to CPLEX for some other, albeit not practical, problem instances where there is ample capacity flexibility. The quality of the solutions produced by our heuristic is dependent on the scenario of products' ordering. In our heuristic, we use the lot-shifting technique to construct this vector. However, this procedure will not always guarantee an ordering scenario that offers the lowest cost, it will only produce a feasible ordering scenario. Having a flexible capacity requires producing sophisticated ordering scenarios to guarantee the lowest cost. Using the lot-shifting technique may not always guarantee producing such ordering scenarios. As a result, our heuristic will produce solutions of a good quality for problems having a tight capacity.

Another important fact about our heuristic is that the time spent to solve problems does not always increases as the size of the problem increases. This can be explained by the examining $While$ statement in **Solve_Model** procedure described in Algorithm 2. We use this procedure to place orders from suppliers to meet demand. The $While$ statement will automatically be exited when demand is met in that period and it will not loop over all suppliers. Therefore, the quicker the demand is met in that period, the fastest the $While$ statement will be executed and therefore a short-time solution is generated. The demand is quickly met when suppliers have a large capacity value. As a result, if the capacity of suppliers is large, our heuristic will produce short-time solutions.

4.5 A procedure for generating feasible instances

In general MIPs are $\mathcal{NP}$_Hard (Schrijver, 1998). For some, even generating a feasible solution may be $\mathcal{NP}$_Complete (Cook, 1971). Therefore, for the problem at hand we are interested first at developing methods for generating feasible solutions and later for an efficient method for finding solutions. In order to test the proposed optimization problem and see the impact of including CSR requirements, in this section we describe a procedure for generating feasible problem instances. The two complicating constraints are the capacity and CSR constraints. We will address each separately in the next two sections.

4.5.1 Capacity constraints

To have a feasible problem, the capacity should meet the demand for all products at all time periods. The following Procedure 1 will allow us to generate instances that satisfy the capacity constraints.

Procedure for generating capacity values:

Step 1:

We will start by specifying the number of products i, number of suppliers j and number of time periods t. After specifying the indices values, we create the following matrix of demand D, where D_{ti} are randomly generated from a Uniform distribution in $[D_{min}, D_{max}]$. We have used $D_{min} = 10$ and $D_{max} = 25$.

$$D = \begin{array}{c c} & \begin{array}{cccc} 1 & 2 & \dots & I \end{array} \\ \begin{array}{c} 1 \\ 2 \\ \vdots \\ T \end{array} & \begin{pmatrix} D_{11} & \vdots & \dots & D_{1I} \\ D_{21} & \vdots & \dots & D_{2I} \\ \vdots & \vdots & \ddots & \vdots \\ D_{T1} & \vdots & \dots & D_{TI} \end{pmatrix} \end{array}$$

Step 2:

Generating the capacity matrix C will be based on the demand matrix values. We start by calculating the cumulative sum of each product's demand for each time period t. Since in our model each supplier has the same capacity in each time period for each product, we will divide each cumulative sum of product's demand by the index of the time period t. For instance, the cumulative sum of products' demand in time period t_n will be divided by n, where $n = 1 \dots T$. The matrix for average cumulative sum for products' demand is:

$$\begin{array}{c c} & \begin{array}{ccccc} 1 & \dots & i & \dots & I \end{array} \\ \begin{array}{c} 1 \\ 2 \\ \vdots \\ T \end{array} & \begin{pmatrix} \frac{D_{11}}{1} & \vdots & \frac{D_{1i}}{1} & \dots & \frac{D_{1I}}{1} \\ \frac{D_{11}+D_{21}}{2} & \vdots & \frac{D_{1i}+D_{2i}}{2} & \dots & \frac{D_{1I}+D_{2I}}{2} \\ \vdots & \vdots & \ddots & \dots & \vdots \\ \frac{D_{11}+\ldots+D_{T1}}{T} & \vdots & \frac{D_{1i}+\ldots+D_{Ti}}{T} & \dots & \frac{D_{1I}+\ldots+D_{TI}}{T} \end{pmatrix} \end{array}$$

Step 3:

After obtaining the average cumulative sum of demand for all products in all time periods, we will choose for each product i the maximum value. After rounding up this value, it will represent the minimum sum of capacities suppliers need to have for product i to have a feasible problem. For instance, for each product i, the sum of all suppliers' capacities should be equal to :

$$SumC_i = \gamma * \lfloor max(\frac{D_{1i}}{1}, \ldots, \frac{D_{1i}+\ldots+D_{Ti}}{T}) \rceil, \quad i = 1 \ldots I, \quad \gamma \geq 1.$$

Where γ is a parameter that we use to allow for capacity flexibility. For example when $\gamma = 1.5$ we are allowing for at least 50% flexibility.

Step 4:

The last step is to obtain the capacity for each supplier and product i. To do so, we divide the capacity's sum $SumC_i$ between all suppliers, where each supplier j will have capacity C_{ji} for product i, such that:

$$\sum_{j=1}^{J} C_{ji} = SumC_i, \quad i = 1 \ldots I.$$

Dividing this the capacity's sum $SumC_i$ between all suppliers is ensured using a multinomial random generator. This generator draws samples from a multinomial distribution. We have used Python's Numpy package **multinomial**[1]generator with number of experiments equal to $sumC_i$ and equal probability values.

To generate the price values P_{ij}, we first generate price values P_i for each product i from a Uniform distribution in $[P_{min}, P_{max}]$. We have used $P_{min} = 30$ and $P_{max} = 50$. For each supplier, the purchase price will be

$$P_{ij} = \omega P_i, \quad \text{where} \quad \omega \quad \text{in} \quad [0.1, 2], \quad i = 1 \ldots I, \quad j = 1 \ldots J.$$

[1] `https://docs.scipy.org/doc/numpy/reference/generated/numpy.random.multinomial.html`

The holding cost values H_i are then calculated as a percentage of the purchase price values:

$$H_i = \beta \frac{\sum_{j=1}^{J} P_{ij}}{J}, \quad \text{where} \quad \beta \quad \text{in} \quad [0.1, 0.2], \quad i = 1 \ldots I.$$

The ordering cost O_j is generated from a Uniform distribution in $[O_{min}, O_{max}]$. We have used $O_{min} = 50$ and $O_{max} = 200$.

In Proposition 1 we prove that our capacity generation procedure leads to feasible capacities constraints.

Proposition 1. *The output of procedure for generating capacity values always results in feasible capacities.*

Proof. Suppose that we order from each supplier j in period t at capacity, i.e., we consider a worst case scenario, so that $Y_{jt} = 1$ and $X_{ijt} = C_j$ for all i, t, and j. It follows that constraint (4.2) becomes $\sum_{k=1}^{t} \sum_{j=1}^{J} C_j \geq \sum_{k=1}^{t} D_{ik}$ for all i and t.
Since we assume that C_j is constant for all t, we get:

$$t \sum_{j=1}^{J} C_j \geq \sum_{k=1}^{t} D_{ik} \quad \text{or} \quad \sum_{j=1}^{J} C_j \geq \frac{1}{t} \sum_{k=1}^{t} D_{ik} \text{ for all } i \text{ and } t.$$

Letting $\sum_{j=1}^{J} C_j = \max\limits_{1 \leq i \leq I} (\frac{1}{t} \sum_{k=1}^{t} D_{ik})$, as in Step 3 of the procedure, will always respect $\sum_{j=1}^{J} C_j \geq \frac{1}{t} \sum_{k=1}^{t} D_{ik}$ for all i and t. □

4.5.2 CSR constraint

To test our optimization model, we generate CSR risk probabilities Pr_j from a Uniform distribution in the range $[0, 1]$. The closer the values to one, the risker the supplier j is. We have chosen the following condition on the CSR risk probabilities.

$$\frac{\sum_{j=1}^{J} Pr_j}{J} \leq \alpha.$$

We note that this condition may not guarantee feasibility of the CSR constraints, however, in all our computational tests it has provided feasible solutions. We can come up with more stricter conditions that can guarantee feasibility. For example, using majorization inequalities theory, we can enforce the above inequality on partial means and add the condition that the sum of orders over all products and periods are majorized over j. This condition is however restrictive and not representative of many practical situations.

4.6 Computational Analysis

Using the feasibility procedures discussed in the previous section, we have generated 216 problem instances. For each of the products, suppliers and periods, we start from a value of 5 and increment it by 10 to a maximum value of 55. The full problem instance sizes can be seen in Appendix A.4.

After generating data instances, we have solved these MIP problems using the solver CPLEX 12.8 in GAMS 25.1. To assess the impact of incorporating CSR requirements in the supplier selection and lot sizing problem, we have solved two version of the MIP problem: one with the CSR constraint and one without that constraint. To study different levels of CSR enforcement, we have varied the CSR target, α, between the values 0.3 (strict target), 0.5 (moderate target), and 0.8 (loose target).

4.6.1 Impact of introducing CSR constraint on cost

In order to study the impact of introducing the CSR constraint on total costs, we show in Table 4.5 a summary of our finding for the 864 problem instances. In this table, we present the cost values' summary for instances solved with the CSR constraint (second column) and those for when CSR is incorporated at α= 0.3, 0.5 and 0.8 in columns 3, 4 and 5, respectively.

	Regular cost	CSR cost		
		$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 0.8$
count	216	216	216	216
mean	191746.66	212732.87	194763.45	191720.70
std. dev.	181840.14	212398.63	186589.25	181824.46
min	4572.33	4975.45	4572.33	4572.33
max	986188.50	1483222.74	1086099.92	986188.50

Table 4.5: Summary of cost values for different problem instances

We can see that as the value of α decreases, i.e., we have a tighter CSR constraint, the average value of the total cost increases.This data can be valuable for a company as they can assess the cost impact of the implementation of a CSR policy. For example, moving from an expected 80% CSR risk to 50% increases the cost by 1.6%. However, moving from 50% to 30% increases the cost by 9.2%. This suggest that the marginal increases in cost is an increasing convex function. It is also interesting to note that the cost variability increases as we increase the CSR requirement. In some supply chain environments this cost uncertainty may add an implicit cost to the total supply chain cost, since, for example, it may make production planning more uncertain. Finally, we note that for problem instances where the total costs are low, the impact of of CSR is not as significant as that for problems where the total costs are large. From Table

4.5 we see that for small cost instances, a stringent CSR requirement increases costs by 8.8% and did not change costs at all for moderate and flexible CSR requirements. The situation is quite different for the large cost instance where a stringent CSR requirement led to an increase of more than 50% in costs. This suggests that CSR implications are more significant for large orders that are more likely to occur with large companies.

To better understand the impact of CSR requirements on costs, we have calculated a ratio of regular costs to instances where CSR requirements are imposed, as follows:

$$ratio_\alpha = 100 * |\frac{cost_{Regular} - cost_\alpha}{cost_{Regular}}|, \quad \text{where } \alpha = \{0.3, 0.5, 0.8\}$$

In Table 4.6 we show the percentages of instances having a cost value that is 3% higher than the regular cost scenario.

$ratio_{0.3}$	$ratio_{0.5}$	$ratio_{0.8}$
76.85%	11.57%	0%

Table 4.6: Percentage of instances having $ratio_\alpha \geq 3\%$.

We see that when we have a tight CSR constraint, 76.85% of the problems have a cost that is 3% higher than the original solution. That percentage decreases to 11.57% for moderate CSR requirements and to 0% for loose requirements. Theses observations confirm that the more we are concerned about maintaining CSR, i.e., we only purchase products from suppliers having the lowest CSR risk probability, the more costly it will be.

In Figure 4.2 we present cost values for all 216 instances for the regular scenario and the moderate CSR requirement. We observe that the behavior of cost values is approximately the same. However, we should note that there could still be some differences in the quantities ordered from suppliers based on their CSR risk probabilities.

We will investigate this issue in the next sections.

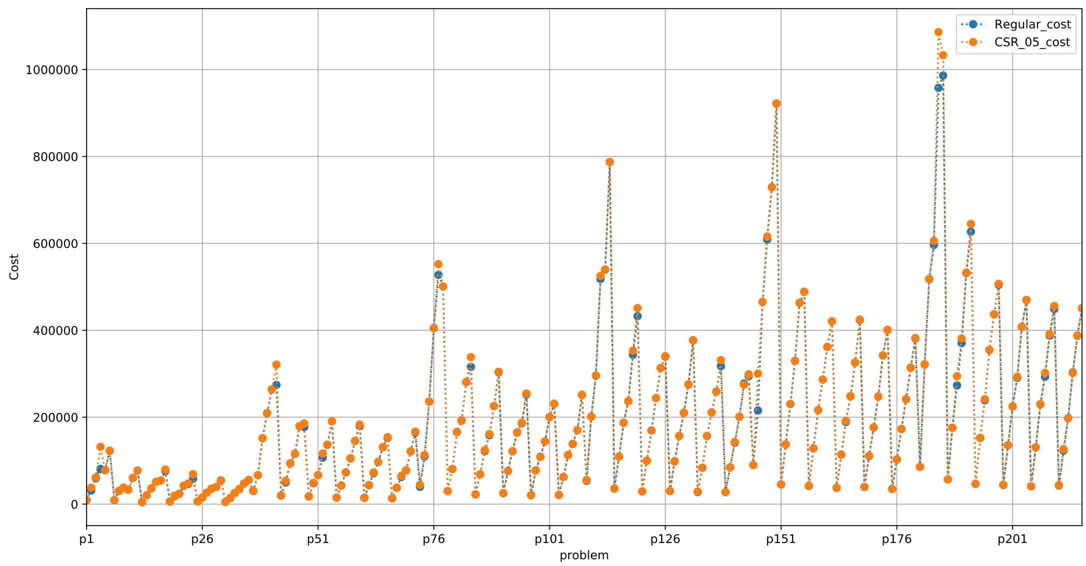


Figure 4.2: Cost comparison of *Regular* and $\alpha = 0.5$ instances

4.6.2 Impact of introducing CSR constraint on number of suppliers

Our goal is to study the impact of CSR enforcement on a company supplier base. One would expect the number of suppliers to be lower problems with a stricter CSR requirement. To test this hypothesis, we have generated problems where the capacity parameter γ can be 1 (low flexibility) or 1.5 (high flexibility) and the CSR constraint parameter α can be 0.3 (strict) or 0.8 (loose). This resulted in 864 $(= 216 \times 4)$ problem instances.

We report on the number of used suppliers and their proportion in Appendix A.2 and Appendix A.3 for $\gamma = 1$ and 1.5 respectively. A summary of the results is shown in Table 4.7.

Capacity parameter γ	1		1.5	
CSR parameter α	0.3	0.8	0.3	0.8
Average supplier ratio	99.11%	99.11%	94.55%	94.60%

Table 4.7: Summary results for impact of including CSR constraint on the number of suppliers

We see from Table 4.7 that the number of suppliers stays almost constant for a given capacity flexibility. Furthermore, an increase in the capacity flexibility by 50% leads to a relatively smaller decrease in the supply base ratio by about 5%.

These results suggest that the supplier base does not change significantly when we enforce CSR. This is indeed a comforting insight for companies that are considering move to CSR and worried that this could disrupt their existing supply network. While the supply base may stay the same, it is quite likely that the ordered quantiles may change significantly. We will investigate this issue in the next section.

4.6.3 Impact of introducing CSR constraint on quantity of products ordered from suppliers

We examine how CSR risk probabilities of suppliers impact the firm ordering pattern. Intuitively, we expected that as the CSR requirements become strict the more will tend to order from suppliers with low CSR risks and the less with suppliers with high CSR risks. To illustrate this statement, we consider the first problem instance where $I = J = T = 5$ and study the quantity of products ordered when suppliers CSR requirement varies in $\{0.3, 0.5, 0.8\}$. The problem instance data is in Table 4.3. The results are presented in Table 4.8.

Supplier	CSR risk probabilities	Regular	CSR		
			$\alpha = 0.8$	$\alpha = 0.5$	$\alpha = 0.3$
j_0	0.67	0	0	0	0
j_1	0.26	22	22	42	248
j_2	0.12	57	57	53	89
j_3	0.60	99	99	99	0
j_4	0.57	249	249	233	90

Table 4.8: Quantity of products ordered per supplier when I=J=T=5.

For $\alpha = 0.8$ there is no change in the order pattern. About 81.50% of the ordered quantities is from suppliers having high CSR risk probabilities, while 18.50% of the order is from suppliers having lower CSR risk probabilities. For $\alpha = 0.5$, we find that 77.75% of the quantities are placed from suppliers having a high CSR risk probabilities, while 22.25% are from suppliers having lower CSR risk probabilities. For $\alpha = 0.3$, about 21.08% of the quantities is from suppliers having a high CSR risk probabilities, while $78,92\%$ is from suppliers having lower CSR risk probabilities. From this analysis we see that the impact of a CSR requirement can be significant in the redistribution of orders among suppliers. For a more strict CSR requirement we note that the supplier base decreased by 25% and the changes in suppliers order ranged from 0% (no change) to more than 1000% (more than tenfold increase). This suggests that firms should plan carefully for the introduction of CSR requirements as that can put heavy pressure on some suppliers, such with supplier 1 in Table 4.8. One way to moderate this effect is to add a constraint that would limit the absolute changes to suppliers orders so that they are within levels that are acceptable to the suppliers.

We conclude by noting that, if desired, our MIP formulation can easily be modified to enforce the exclusion of some suppliers, such as based on a minimum required Pr_j, and to enforce a minimum supplier order quantity.

4.7 Conclusion

We presented a supplier selection and lot-sizing model that incorporates a CSR requirement. We consider multiple products, suppliers and periods as well as supplier capacities. We show that the problem is NP-hard and propose procedures for generating feasible instances as well as a heuristic to solve the problem. We have showed, computationally, the impact of incorporating CSR requirements on costs, suppliers' orders and capacities.

Chapter 5

Conclusion and discussion

Monitoring CSR in a global supply chain is a challenging process. We presented a practical model for predicting supplier deviance behaviour. We text mine data for news sources to look for signals of abnormal supplier activities and we build a predictive model using machine learning techniques in order to quantify the CSR risk measure of a given supplier. To do this, we have represented risks related to suppliers by two classes: labour issues and health and safety issues. We also have added a third class to represent suppliers who respect social responsibility. Based on this partition, we have collected newspapers from Factiva global news database using its search tool and we have used classification algorithms to classify the retrieved data into one of the three classes. The output of this predictive model was the CSR risk probability representing the sum of classification probabilities for the labor and health and safety classes.

The CSR risk probabilities were used to build a supplier selection and lot-sizing model that incorporates CSR. multiple products, suppliers and periods and suppliers' capacities. In order to study our model, we have implemented a method for generating feasible problem instances. After producing several instances, we have studied the impact of introducing CSR on aspects such as cost, quantity of products ordered and suppliers' capacity. We found that the stricter the CSR constraint is the more it is going to cost, but having an average CSR constraint will not increase significantly the cost and will help respect CSR on average. Studying the second aspect, we have found that the products' orders placed depends also on the value of the CSR constraint. We noticed that as the value of α decreases, i.e., having a stricter CSR constraint, the percentage of orders placed from suppliers with low risk increases. Moreover, we found that CSR affects suppliers' capacity. In addition, we have developed a heuristic procedure to solve the problem. Our computational results show that our heuristic produces good quality solutions. The solutions produced by our algorithm remain computationally expensive compared to CPLEX. However, short-time solutions can be produced if the suppliers' capacities are high. Moreover, our heuristic will produce good solutions for problems having a tight capacity.

Our work can be extended in several ways. First, when building the predictive model, we may add other data sources, such as internal documents of the company or emails and investigation reports. Another aspect that can also be improved is testing other self-learning algorithms to perform the classification process in order to attain higher accuracy levels. A third venue for future research is to develop hybrid algorithms that mix traditional decomposition methods, such Benders or Lagrangian, with machine learning. Future work may consider improving the heuristic by developing a procedure that produces a procurement scenario that generates the lowest cost.

Another line of research could confider the long term supplier selection problem. In these problems we address questions of how often should we solve the problem or how many suppliers should we have. In the present work, we dealt with the selection of the best subset of suppliers among the existing alternatives. This assumes that the number of suppliers to be selected and time periods are already determined. However, in real-world cases, the determination of the number of suppliers, the relationship mode with them and time periods are chosen according to the characteristics of the firm, the product and the market.

Appendix A

Appendix

A.1 Apple and Foxconn case results

Column names description:

- n : news article number
- Date: news article date of publication
- Actual: actual class of news article (-1): negative examples, (0): labour issues class, (1): health and safety issues class
- x: probability news article belongs to labour issues class
- y: probability news article belongs to health and safety issues class
- z: probability news article belongs to no issues class
- Predicted: predicted class of news article (-1): negative examples, (0): labour issues class, (1): health and safety issues class
- r_n: CSR risk probability per document

n	**Date**	**Actual**	x	y	z	**Predicted**	r_n
1,	6/15/2017,	-1	0.00	0.00	1.00	-1	0.00
2,	6/7/2010,	0	0.88	0.00	0.12	0	0.88
3,	3/30/2012,	0	0.66	0.34	0.00	0	1.00
4,	3/9/2017,	-1	0.00	0.00	1.00	-1	0.00
5,	7/14/2015,	0	1.00	0.00	0.00	0	1.00
6,	7/22/2009,	1	0.61	0.37	0.01	0	0.99
7,	8/18/2017,	-1	0.00	0.00	1.00	-1	0.00
8,	8/7/2017,	-1	0.20	0.00	0.80	-1	0.20
9,	5/31/2013,	0	0.95	0.05	0.00	0	1.00
10,	5/21/2010,	1	0.12	0.88	0.00	1	1.00
11,	8/2/2010,	0	0.87	0.13	0.00	0	1.00
12,	6/12/2010,	1	0.24	0.76	0.00	1	1.00
13,	10/9/2014,	0	0.50	0.00	0.50	0	0.50
14,	10/9/2017,	-1	0.00	0.00	1.00	-1	0.00
15,	9/24/2013,	1	0.00	0.90	0.10	1	0.90
16,	5/4/2018,	-1	0.00	0.02	0.98	-1	0.02
17,	3/28/2018,	-1	0.18	0.00	0.82	-1	0.18
18,	9/5/2017,	-1	0.00	0.00	1.00	-1	0.00
19,	4/1/2012,	0	0.85	0.00	0.15	0	0.85
20,	10/6/2012,	0	0.94	0.00	0.06	0	0.94
21,	5/23/2018,	-1	0.05	0.00	0.95	-1	0.05
22,	5/15/2014,	0	0.94	0.01	0.05	0	0.95
23,	6/8/2010,	0	1.00	0.00	0.00	0	1.00
24,	3/2/2014,	0	0.52	0.00	0.48	0	0.52
25,	1/21/2013,	0	0.96	0.04	0.00	0	1.00
26,	8/17/2013,	0	0.26	0.45	0.29	1	0.71
27,	2/5/2013,	0	0.86	0.14	0.00	0	1.00
28,	6/8/2010,	0	1.00	0.00	0.00	0	1.00
29,	3/28/2017,	-1	0.32	0.00	0.68	-1	0.32
30,	6/2/2010,	0	0.34	0.37	0.29	1	0.71
31,	2/4/2013,	0	0.91	0.09	0.00	0	1.00
32,	4/12/2018,	-1	0.00	0.00	1.00	-1	0.00
33,	5/28/2010,	0	0.44	0.56	0.00	1	1.00
34,	11/26/2017,	0	0.54	0.00	0.46	0	0.54
35,	6/2/2010,	0	1.00	0.00	0.00	0	1.00
36,	6/23/2017,	-1	0.00	0.00	1.00	-1	0.00
37,	5/10/2017,	-1	0.00	0.00	1.00	-1	0.00
38,	10/6/2012,	0	1.00	0.00	0.00	0	1.00
39,	5/27/2010,	1	0.00	0.51	0.49	1	0.51
40,	1/13/2012,	0	0.83	0.17	0.00	0	1.00
41,	5/13/2010,	1	0.00	1.00	0.00	1	1.00
42,	10/24/2017,	-1	0.00	0.00	1.00	-1	0.00
43,	10/10/2014,	0	0.65	0.00	0.35	0	0.65
44,	3/2/2017,	-1	0.00	0.00	1.00	-1	0.00
45,	3/23/2018,	-1	0.00	0.00	1.00	-1	0.00
46,	7/28/2017,	-1	0.45	0.00	0.55	-1	0.45

47,	2/4/2013,	0	1.00	0.00	0.00	0	1.00
48,	4/3/2018,	-1	0.00	0.00	1.00	-1	0.00
49,	8/10/2017,	-1	0.00	0.00	1.00	-1	0.00
50,	5/25/2017,	-1	0.00	0.06	0.94	-1	0.06
51,	3/20/2017,	-1	0.41	0.00	0.59	-1	0.41
52,	9/25/2013,	1	0.00	0.65	0.35	1	0.65
———,	———,	———	———	———	———	Average	0.56

A.2 Number of suppliers for $\gamma = 1$

problem	**Actual # suppliers**	$\#nbsuppliers_{\alpha=0.3}$	$\#suppliers_{\alpha=0.8}$	$ratio_{\alpha=0.3}$	$ratio_{\alpha=0.8}$
p1	5	5	5	1.00	1.00
p2	5	5	5	1.00	1.00
p3	5	5	5	1.00	1.00
p4	5	5	5	1.00	1.00
p5	5	5	5	1.00	1.00
p6	5	5	5	1.00	1.00
p7	15	15	15	1.00	1.00
p8	15	15	15	1.00	1.00
p9	15	15	15	1.00	1.00
p10	15	15	15	1.00	1.00
p11	15	15	15	1.00	1.00
p12	15	15	15	1.00	1.00
p13	25	24	24	0.96	0.96
p14	25	25	25	1.00	1.00
p15	25	25	25	1.00	1.00
p16	25	25	25	1.00	1.00
p17	25	24	24	0.96	0.96
p18	25	25	25	1.00	1.00
p19	35	35	35	1.00	1.00
p20	35	32	32	0.91	0.91
p21	35	32	32	0.91	0.91
p22	35	31	31	0.89	0.89
p23	35	34	34	0.97	0.97
p24	35	33	33	0.94	0.94
p25	45	40	40	0.89	0.89
p26	45	40	40	0.89	0.89
p27	45	41	41	0.91	0.91
p28	45	39	39	0.87	0.87

p29	45	42	42	0.93	0.93
p30	45	40	40	0.89	0.89
p31	55	48	48	0.87	0.87
p32	55	44	44	0.80	0.80
p33	55	49	49	0.89	0.89
p34	55	49	49	0.89	0.89
p35	55	48	48	0.87	0.87
p36	55	46	46	0.84	0.84
p37	5	5	5	1.00	1.00
p38	5	5	5	1.00	1.00
p39	5	5	5	1.00	1.00
p40	5	5	5	1.00	1.00
p41	5	5	5	1.00	1.00
p42	5	5	5	1.00	1.00
p43	15	15	15	1.00	1.00
p44	15	15	15	1.00	1.00
p45	15	15	15	1.00	1.00
p46	15	15	15	1.00	1.00
p47	15	15	15	1.00	1.00
p48	15	15	15	1.00	1.00
p49	25	25	25	1.00	1.00
p50	25	25	25	1.00	1.00
p51	25	25	25	1.00	1.00
p52	25	25	25	1.00	1.00
p53	25	25	25	1.00	1.00
p54	25	25	25	1.00	1.00
p55	35	35	35	1.00	1.00
p56	35	35	35	1.00	1.00
p57	35	35	35	1.00	1.00
p58	35	35	35	1.00	1.00
p59	35	35	35	1.00	1.00
p60	35	35	35	1.00	1.00
p61	45	45	45	1.00	1.00
p62	45	45	45	1.00	1.00
p63	45	45	45	1.00	1.00
p64	45	45	45	1.00	1.00
p65	45	45	45	1.00	1.00
p66	45	45	45	1.00	1.00
p67	55	55	55	1.00	1.00
p68	55	55	55	1.00	1.00
p69	55	55	55	1.00	1.00
p70	55	55	55	1.00	1.00
p71	55	55	55	1.00	1.00
p72	55	55	55	1.00	1.00
p73	5	5	5	1.00	1.00
p74	5	5	5	1.00	1.00
p75	5	5	5	1.00	1.00

p76	5	5	5	1.00	1.00
p77	5	5	5	1.00	1.00
p78	5	5	5	1.00	1.00
p79	15	15	15	1.00	1.00
p80	15	15	15	1.00	1.00
p81	15	15	15	1.00	1.00
p82	15	15	15	1.00	1.00
p83	15	15	15	1.00	1.00
p84	15	15	15	1.00	1.00
p85	25	25	25	1.00	1.00
p86	25	25	25	1.00	1.00
p87	25	25	25	1.00	1.00
p88	25	25	25	1.00	1.00
p89	25	25	25	1.00	1.00
p90	25	25	25	1.00	1.00
p91	35	35	35	1.00	1.00
p92	35	35	35	1.00	1.00
p93	35	35	35	1.00	1.00
p94	35	35	35	1.00	1.00
p95	35	35	35	1.00	1.00
p96	35	35	35	1.00	1.00
p97	45	45	45	1.00	1.00
p98	45	45	45	1.00	1.00
p99	45	45	45	1.00	1.00
p100	45	45	45	1.00	1.00
p101	45	45	45	1.00	1.00
p102	45	45	45	1.00	1.00
p103	55	55	55	1.00	1.00
p104	55	55	55	1.00	1.00
p105	55	55	55	1.00	1.00
p106	55	55	55	1.00	1.00
p107	55	55	55	1.00	1.00
p108	55	55	55	1.00	1.00
p109	5	5	5	1.00	1.00
p110	5	5	5	1.00	1.00
p111	5	5	5	1.00	1.00
p112	5	5	5	1.00	1.00
p113	5	5	5	1.00	1.00
p114	5	5	5	1.00	1.00
p115	15	15	15	1.00	1.00
p116	15	15	15	1.00	1.00
p117	15	15	15	1.00	1.00
p118	15	15	15	1.00	1.00
p119	15	15	15	1.00	1.00
p120	15	15	15	1.00	1.00
p121	25	25	25	1.00	1.00
p122	25	25	25	1.00	1.00

p123	25	25	25	1.00	1.00
p124	25	25	25	1.00	1.00
p125	25	25	25	1.00	1.00
p126	25	25	25	1.00	1.00
p127	35	35	35	1.00	1.00
p128	35	35	35	1.00	1.00
p129	35	35	35	1.00	1.00
p130	35	35	35	1.00	1.00
p131	35	35	35	1.00	1.00
p132	35	35	35	1.00	1.00
p133	45	45	45	1.00	1.00
p134	45	45	45	1.00	1.00
p135	45	45	45	1.00	1.00
p136	45	45	45	1.00	1.00
p137	45	45	45	1.00	1.00
p138	45	45	45	1.00	1.00
p139	55	55	55	1.00	1.00
p140	55	55	55	1.00	1.00
p141	55	55	55	1.00	1.00
p142	55	55	55	1.00	1.00
p143	55	55	55	1.00	1.00
p144	55	55	55	1.00	1.00
p145	5	5	5	1.00	1.00
p146	5	5	5	1.00	1.00
p147	5	5	5	1.00	1.00
p148	5	5	5	1.00	1.00
p149	5	5	5	1.00	1.00
p150	5	5	5	1.00	1.00
p151	15	15	15	1.00	1.00
p152	15	15	15	1.00	1.00
p153	15	15	15	1.00	1.00
p154	15	15	15	1.00	1.00
p155	15	15	15	1.00	1.00
p156	15	15	15	1.00	1.00
p157	25	25	25	1.00	1.00
p158	25	25	25	1.00	1.00
p159	25	25	25	1.00	1.00
p160	25	25	25	1.00	1.00
p161	25	25	25	1.00	1.00
p162	25	25	25	1.00	1.00
p163	35	35	35	1.00	1.00
p164	35	35	35	1.00	1.00
p165	35	35	35	1.00	1.00
p166	35	35	35	1.00	1.00
p167	35	35	35	1.00	1.00
p168	35	35	35	1.00	1.00
p169	45	45	45	1.00	1.00

p170	45	45	45	1.00	1.00
p171	45	45	45	1.00	1.00
p172	45	45	45	1.00	1.00
p173	45	45	45	1.00	1.00
p174	45	45	45	1.00	1.00
p175	55	55	55	1.00	1.00
p176	55	55	55	1.00	1.00
p177	55	55	55	1.00	1.00
p178	55	55	55	1.00	1.00
p179	55	55	55	1.00	1.00
p180	55	55	55	1.00	1.00
p181	5	5	5	1.00	1.00
p182	5	5	5	1.00	1.00
p183	5	5	5	1.00	1.00
p184	5	5	5	1.00	1.00
p185	5	5	5	1.00	1.00
p186	5	5	5	1.00	1.00
p187	15	15	15	1.00	1.00
p188	15	15	15	1.00	1.00
p189	15	15	15	1.00	1.00
p190	15	15	15	1.00	1.00
p191	15	15	15	1.00	1.00
p192	15	15	15	1.00	1.00
p193	25	25	25	1.00	1.00
p194	25	25	25	1.00	1.00
p195	25	25	25	1.00	1.00
p196	25	25	25	1.00	1.00
p197	25	25	25	1.00	1.00
p198	25	25	25	1.00	1.00
p199	35	35	35	1.00	1.00
p200	35	35	35	1.00	1.00
p201	35	35	35	1.00	1.00
p202	35	35	35	1.00	1.00
p203	35	35	35	1.00	1.00
p204	35	35	35	1.00	1.00
p205	45	45	45	1.00	1.00
p206	45	45	45	1.00	1.00
p207	45	45	45	1.00	1.00
p208	45	45	45	1.00	1.00
p209	45	45	45	1.00	1.00
p210	45	45	45	1.00	1.00
p211	55	55	55	1.00	1.00
p212	55	55	55	1.00	1.00
p213	55	55	55	1.00	1.00
p214	55	55	55	1.00	1.00
p215	55	55	55	1.00	1.00
p216	55	55	55	1.00	1.00

————	Average	29.58796296	29.58796296	0.99	0.99
————	Standard deviation	16.74186074	16.74186074	0.03	0.03

A.3 Number of suppliers for $\gamma = 1.5$

problem	**Actual# suppliers**	$\#nbsuppliers_{\alpha=0.3}$	$\#suppliers_{\alpha=0.8}$	$ratio_{\alpha=0.3}$	$ratio_{\alpha=0.8}$
p1	5	5	5	1.00	1.00
p2	5	5	5	1.00	1.00
p3	5	5	5	1.00	1.00
p4	5	5	5	1.00	1.00
p5	5	5	5	1.00	1.00
p6	5	5	5	1.00	1.00
p7	15	12	12	0.80	0.80
p8	15	14	14	0.93	0.93
p9	15	15	15	1.00	1.00
p10	15	13	13	0.87	0.87
p11	15	13	13	0.87	0.87
p12	15	14	15	0.93	1.00
p13	25	17	17	0.68	0.68
p14	25	17	17	0.68	0.68
p15	25	18	18	0.72	0.72
p16	25	19	19	0.76	0.76
p17	25	21	21	0.84	0.84
p18	25	23	23	0.92	0.92
p19	35	22	22	0.63	0.63
p20	35	23	23	0.66	0.66
p21	35	25	25	0.71	0.71
p22	35	27	27	0.77	0.77
p23	35	25	25	0.71	0.71
p24	35	27	27	0.77	0.77
p25	45	26	26	0.58	0.58
p26	45	27	27	0.60	0.60
p27	45	30	29	0.67	0.64
p28	45	32	32	0.71	0.71
p29	45	32	32	0.71	0.71
p30	45	33	34	0.73	0.76
p31	55	29	29	0.53	0.53
p32	55	30	32	0.55	0.58
p33	55	38	35	0.69	0.64

p34	55	30	30	0.55	0.55
p35	55	40	40	0.73	0.73
p36	55	37	39	0.67	0.71
p37	5	5	5	1.00	1.00
p38	5	5	5	1.00	1.00
p39	5	5	5	1.00	1.00
p40	5	5	5	1.00	1.00
p41	5	5	5	1.00	1.00
p42	5	5	5	1.00	1.00
p43	15	15	15	1.00	1.00
p44	15	15	15	1.00	1.00
p45	15	15	15	1.00	1.00
p46	15	15	15	1.00	1.00
p47	15	15	15	1.00	1.00
p48	15	15	15	1.00	1.00
p49	25	23	23	0.92	0.92
p50	25	24	24	0.96	0.96
p51	25	24	24	0.96	0.96
p52	25	23	23	0.92	0.92
p53	25	24	24	0.96	0.96
p54	25	24	24	0.96	0.96
p55	35	31	31	0.89	0.89
p56	35	33	33	0.94	0.94
p57	35	33	32	0.94	0.91
p58	35	32	32	0.91	0.91
p59	35	33	33	0.94	0.94
p60	35	33	33	0.94	0.94
p61	45	35	35	0.78	0.78
p62	45	41	42	0.91	0.93
p63	45	38	38	0.84	0.84
p64	45	39	39	0.87	0.87
p65	45	39	39	0.87	0.87
p66	45	41	41	0.91	0.91
p67	55	39	39	0.71	0.71
p68	55	43	43	0.78	0.78
p69	55	44	43	0.80	0.78
p70	55	46	46	0.84	0.84
p71	55	48	47	0.87	0.85
p72	55	43	43	0.78	0.78
p73	5	5	5	1.00	1.00
p74	5	5	5	1.00	1.00
p75	5	5	5	1.00	1.00
p76	5	5	5	1.00	1.00
p77	5	5	5	1.00	1.00
p78	5	5	5	1.00	1.00
p79	15	15	15	1.00	1.00
p80	15	15	15	1.00	1.00

p81	15	15	15	1.00	1.00
p82	15	15	15	1.00	1.00
p83	15	15	15	1.00	1.00
p84	15	15	15	1.00	1.00
p85	25	25	25	1.00	1.00
p86	25	25	25	1.00	1.00
p87	25	24	25	0.96	1.00
p88	25	25	25	1.00	1.00
p89	25	25	25	1.00	1.00
p90	25	25	25	1.00	1.00
p91	35	35	35	1.00	1.00
p92	35	35	34	1.00	0.97
p93	35	35	35	1.00	1.00
p94	35	35	35	1.00	1.00
p95	35	35	35	1.00	1.00
p96	35	35	35	1.00	1.00
p97	45	44	44	0.98	0.98
p98	45	41	41	0.91	0.91
p99	45	43	43	0.96	0.96
p100	45	43	43	0.96	0.96
p101	45	45	45	1.00	1.00
p102	45	44	41	0.98	0.91
p103	55	51	51	0.93	0.93
p104	55	48	51	0.87	0.93
p105	55	52	52	0.95	0.95
p106	55	54	54	0.98	0.98
p107	55	51	51	0.93	0.93
p108	55	50	50	0.91	0.91
p109	5	5	5	1.00	1.00
p110	5	5	5	1.00	1.00
p111	5	5	5	1.00	1.00
p112	5	5	5	1.00	1.00
p113	5	5	5	1.00	1.00
p114	5	5	5	1.00	1.00
p115	15	15	15	1.00	1.00
p116	15	15	15	1.00	1.00
p117	15	15	15	1.00	1.00
p118	15	15	15	1.00	1.00
p119	15	15	15	1.00	1.00
p120	15	15	15	1.00	1.00
p121	25	25	25	1.00	1.00
p122	25	25	25	1.00	1.00
p123	25	25	25	1.00	1.00
p124	25	25	25	1.00	1.00
p125	25	25	25	1.00	1.00
p126	25	25	25	1.00	1.00
p127	35	35	35	1.00	1.00

p128	35	35	35	1.00	1.00
p129	35	35	35	1.00	1.00
p130	35	35	35	1.00	1.00
p131	35	35	35	1.00	1.00
p132	35	35	35	1.00	1.00
p133	45	44	44	0.98	0.98
p134	45	45	45	1.00	1.00
p135	45	45	45	1.00	1.00
p136	45	45	45	1.00	1.00
p137	45	43	44	0.96	0.98
p138	45	45	45	1.00	1.00
p139	55	53	54	0.96	0.98
p140	55	52	53	0.95	0.96
p141	55	54	54	0.98	0.98
p142	55	54	54	0.98	0.98
p143	55	55	55	1.00	1.00
p144	55	55	55	1.00	1.00
p145	5	5	5	1.00	1.00
p146	5	5	5	1.00	1.00
p147	5	5	5	1.00	1.00
p148	5	5	5	1.00	1.00
p149	5	5	5	1.00	1.00
p150	5	5	5	1.00	1.00
p151	15	15	15	1.00	1.00
p152	15	15	15	1.00	1.00
p153	15	15	15	1.00	1.00
p154	15	15	15	1.00	1.00
p155	15	15	15	1.00	1.00
p156	15	15	15	1.00	1.00
p157	25	25	25	1.00	1.00
p158	25	25	25	1.00	1.00
p159	25	25	25	1.00	1.00
p160	25	25	25	1.00	1.00
p161	25	25	25	1.00	1.00
p162	25	25	25	1.00	1.00
p163	35	35	35	1.00	1.00
p164	35	35	35	1.00	1.00
p165	35	35	35	1.00	1.00
p166	35	35	35	1.00	1.00
p167	35	35	35	1.00	1.00
p168	35	35	35	1.00	1.00
p169	45	45	45	1.00	1.00
p170	45	45	45	1.00	1.00
p171	45	45	45	1.00	1.00
p172	45	45	45	1.00	1.00
p173	45	45	45	1.00	1.00
p174	45	45	45	1.00	1.00

p175	55	54	54	0.98	0.98
p176	55	55	55	1.00	1.00
p177	55	55	55	1.00	1.00
p178	55	55	55	1.00	1.00
p179	55	55	55	1.00	1.00
p180	55	55	55	1.00	1.00
p181	5	5	5	1.00	1.00
p182	5	5	5	1.00	1.00
p183	5	5	5	1.00	1.00
p184	5	5	5	1.00	1.00
p185	5	5	5	1.00	1.00
p186	5	5	5	1.00	1.00
p187	15	15	15	1.00	1.00
p188	15	15	15	1.00	1.00
p189	15	15	15	1.00	1.00
p190	15	15	15	1.00	1.00
p191	15	15	15	1.00	1.00
p192	15	15	15	1.00	1.00
p193	25	25	25	1.00	1.00
p194	25	25	25	1.00	1.00
p195	25	25	25	1.00	1.00
p196	25	25	25	1.00	1.00
p197	25	25	25	1.00	1.00
p198	25	25	25	1.00	1.00
p199	35	35	35	1.00	1.00
p200	35	35	35	1.00	1.00
p201	35	35	35	1.00	1.00
p202	35	35	35	1.00	1.00
p203	35	35	35	1.00	1.00
p204	35	35	35	1.00	1.00
p205	45	45	45	1.00	1.00
p206	45	45	45	1.00	1.00
p207	45	45	45	1.00	1.00
p208	45	45	45	1.00	1.00
p209	45	45	45	1.00	1.00
p210	45	45	45	1.00	1.00
p211	55	55	55	1.00	1.00
p212	55	54	54	0.98	0.98
p213	55	55	55	1.00	1.00
p214	55	55	55	1.00	1.00
p215	55	55	55	1.00	1.00
p216	55	55	55	1.00	1.00
————	Average	27.70833333	27.72222222	0.95	0.95
————	Standard deviation	15.6458814	15.65846726	0.11	0.11

A.4 Heuristic and CPLEX summary results

problem	I	J	T	$cost_{Heuristic}$	$cost_{Cplex}$	**ratio**	$time_{Heuristic}$	$time_{Cplex}$
p1	5	5	5	22397.08	22397.08	0.0000	1.24	0.18
p2	5	5	15	57058.54	56565.58	0.0087	4.22	0.22
p3	5	5	25	95932.11	93465.81	0.0264	11.25	0.36
p4	5	5	35	143780.40	139412.76	0.0313	14.76	0.40
p5	5	5	45	182456.88	180858.87	0.0088	11.39	0.54
p6	5	5	55	199670.77	198411.43	0.0063	12.72	0.74
p7	5	15	5	26081.96	25731.40	0.0136	3.01	0.23
p8	5	15	15	77538.71	77109.51	0.0056	10.41	0.40
p9	5	15	25	130254.45	125993.69	0.0338	12.19	0.65
p10	5	15	35	196340.84	189748.45	0.0347	20.97	1.24
p11	5	15	45	282068.89	271271.29	0.0398	34.68	2.30
p12	5	15	55	331720.03	325402.94	0.0194	31.54	2.38
p13	5	25	5	28864.05	28068.79	0.0283	9.45	0.25
p14	5	25	15	98718.73	96261.49	0.0255	12.20	0.49
p15	5	25	25	177893.80	169802.32	0.0477	39.13	0.96
p16	5	25	35	236402.94	225079.63	0.0503	29.50	2.59
p17	5	25	45	274845.98	263649.13	0.0425	43.15	2.05
p18	5	25	55	343433.35	319076.69	0.0763	54.31	4.16
p19	5	35	5	35462.60	30605.15	0.1587	10.49	0.34
p20	5	35	15	105584.20	100871.52	0.0467	34.37	0.90
p21	5	35	25	177921.62	164343.79	0.0826	43.40	1.07
p22	5	35	35	248264.63	226077.09	0.0981	53.89	3.50
p23	5	35	45	331858.10	299706.43	0.1073	57.72	4.91
p24	5	35	55	425005.94	388552.46	0.0938	118.84	4.34
p25	5	45	5	40899.26	38152.22	0.0720	8.38	0.35
p26	5	45	15	124290.75	112048.01	0.1093	27.71	0.82
p27	5	45	25	202445.86	176777.16	0.1452	56.74	1.42
p28	5	45	35	289413.30	267129.22	0.0834	76.16	2.32
p29	5	45	45	376636.01	342766.28	0.0988	84.31	2.42
p30	5	45	55	426061.68	351037.33	0.2137	81.66	11.21
p31	5	55	5	52215.84	47957.64	0.0888	10.15	0.56
p32	5	55	15	136810.36	128996.47	0.0606	28.33	1.09
p33	5	55	25	212254.29	183451.76	0.1570	78.16	2.45
p34	5	55	35	304568.11	247308.15	0.2315	79.77	2.78
p35	5	55	45	411225.75	372492.96	0.1040	102.20	4.67
p36	5	55	55	485302.41	444484.93	0.0918	109.19	6.12
p37	15	5	5	56887.26	56840.02	0.0008	3.41	0.28
p38	15	5	15	153294.70	152334.03	0.0063	11.56	0.48
p39	15	5	25	304207.20	301719.18	0.0082	21.35	0.66
p40	15	5	35	392233.55	389057.85	0.0082	25.78	1.09
p41	15	5	45	545139.43	539528.49	0.0104	31.27	1.13
p42	15	5	55	707171.87	703554.73	0.0051	47.50	1.81

p43	15	15	5	59323.26	57938.28	0.0239	10.20	0.49
p44	15	15	15	182482.41	175508.10	0.0397	56.05	0.96
p45	15	15	25	339003.63	325139.78	0.0426	56.76	1.94
p46	15	15	35	471910.30	467355.07	0.0097	52.32	2.28
p47	15	15	45	550984.03	518463.77	0.0627	510.14	3.57
p48	15	15	55	642943.71	636719.22	0.0098	67.61	4.35
p49	15	25	5	70985.80	68473.02	0.0367	16.65	0.51
p50	15	25	15	188662.53	187180.40	0.0079	28.68	1.39
p51	15	25	25	336442.37	333608.72	0.0085	46.35	2.43
p52	15	25	35	476278.35	472960.31	0.0070	67.58	3.85
p53	15	25	45	647270.49	643816.10	0.0054	88.76	4.78
p54	15	25	55	695854.88	689654.01	0.0090	113.82	7.53
p55	15	35	5	72863.58	70350.35	0.0357	29.35	0.71
p56	15	35	15	212017.04	210374.00	0.0078	43.79	1.67
p57	15	35	25	381366.26	377181.02	0.0111	65.63	3.13
p58	15	35	35	543127.61	517422.69	0.0497	207.05	5.31
p59	15	35	45	660349.35	645718.62	0.0227	114.60	8.34
p60	15	35	55	809308.39	787901.75	0.0272	142.76	10.25
p61	15	45	5	79058.66	77706.68	0.0174	17.65	0.85
p62	15	45	15	258366.73	246368.05	0.0487	99.35	2.45
p63	15	45	25	418242.45	404978.78	0.0328	93.54	3.87
p64	15	45	35	536730.72	521165.30	0.0299	120.13	5.59
p65	15	45	45	779684.83	745592.84	0.0457	236.90	8.16
p66	15	45	55	935482.36	920748.15	0.0160	190.38	10.21
p67	15	55	5	82022.59	79856.97	0.0271	33.10	0.91
p68	15	55	15	264744.34	259601.28	0.0198	64.11	2.64
p69	15	55	25	422434.80	407358.34	0.0370	130.74	4.94
p70	15	55	35	612045.49	590376.73	0.0367	204.89	7.18
p71	15	55	45	827303.49	792151.41	0.0444	368.78	10.35
p72	15	55	55	923283.47	897918.58	0.0282	273.42	15.20
p73	25	5	5	101874.77	101283.67	0.0058	8.03	0.48
p74	25	5	15	275021.11	274200.91	0.0030	19.63	0.73
p75	25	5	25	419575.80	416335.32	0.0078	28.06	1.08
p76	25	5	35	651199.50	646724.76	0.0069	40.75	1.48
p77	25	5	45	810559.29	805838.94	0.0059	53.66	2.12
p78	25	5	55	1090193.30	1083665.90	0.0060	73.28	3.04
p79	25	15	5	89614.60	89229.95	0.0043	13.78	0.83
p80	25	15	15	294967.75	281133.01	0.0492	66.70	1.95
p81	25	15	25	501219.27	497457.15	0.0076	65.42	2.41
p82	25	15	35	636267.09	631139.21	0.0081	83.71	3.83
p83	25	15	45	847776.04	844332.86	0.0041	104.68	5.01
p84	25	15	55	1134605.93	1119369.11	0.0136	151.11	7.80
p85	25	25	5	104183.43	103955.49	0.0022	18.43	0.87
p86	25	25	15	329328.74	318844.36	0.0329	427.74	2.18
p87	25	25	25	507056.12	503911.85	0.0062	101.47	3.80
p88	25	25	35	719597.67	716578.03	0.0042	125.84	5.72
p89	25	25	45	910239.47	904593.13	0.0062	163.11	7.45

p90	25	25	55	1239714.50	1236100.27	0.0029	212.51	9.75
p91	25	35	5	111017.06	108807.84	0.0203	51.80	1.07
p92	25	35	15	334464.29	325694.40	0.0269	85.08	2.72
p93	25	35	25	577082.19	558408.50	0.0334	239.26	5.50
p94	25	35	35	691526.49	685351.33	0.0090	223.29	6.86
p95	25	35	45	960696.66	956856.02	0.0040	257.13	11.84
p96	25	35	55	1208439.81	1202411.50	0.0050	345.15	15.90
p97	25	45	5	117906.13	113893.18	0.0352	90.81	1.22
p98	25	45	15	334603.91	330845.97	0.0114	103.03	3.51
p99	25	45	25	577063.14	570542.47	0.0114	200.83	7.10
p100	25	45	35	774954.82	767978.37	0.0091	275.39	9.72
p101	25	45	45	1019607.15	1013945.91	0.0056	364.13	11.21
p102	25	45	55	1289964.27	1261379.91	0.0227	547.47	21.75
p103	25	55	5	125117.36	121162.98	0.0326	77.19	1.73
p104	25	55	15	384669.39	375426.32	0.0246	961.38	3.90
p105	25	55	25	562432.23	554023.96	0.0152	224.39	6.47
p106	25	55	35	855387.56	848896.97	0.0076	336.08	12.25
p107	25	55	45	1107938.26	1081036.75	0.0249	422.66	21.83
p108	25	55	55	1330432.58	1313396.93	0.0130	480.79	24.93
p109	35	5	5	115125.19	114651.04	0.0041	12.73	0.62
p110	35	5	15	369371.65	367319.28	0.0056	23.83	0.98
p111	35	5	25	666080.76	652484.07	0.0208	53.21	1.51
p112	35	5	35	820849.72	797039.92	0.0299	89.20	2.31
p113	35	5	45	1110547.28	1104666.24	0.0053	70.86	2.96
p114	35	5	55	1243262.82	1234807.89	0.0068	98.38	3.90
p115	35	15	5	136435.99	132769.21	0.0276	32.72	0.89
p116	35	15	15	422606.07	404242.91	0.0454	143.21	1.94
p117	35	15	25	694884.69	690875.41	0.0058	111.84	3.02
p118	35	15	35	881497.91	874426.30	0.0081	120.82	5.29
p119	35	15	45	1096128.31	1087791.15	0.0077	155.57	7.77
p120	35	15	55	1552745.43	1469647.47	0.0565	421.30	9.66
p121	35	25	5	141232.03	137881.26	0.0243	68.68	1.66
p122	35	25	15	422030.17	409838.19	0.0297	131.62	2.90
p123	35	25	25	673933.93	671111.23	0.0042	160.28	5.70
p124	35	25	35	997021.00	960243.87	0.0383	352.08	8.65
p125	35	25	45	1210344.70	1156682.62	0.0464	437.53	10.66
p126	35	25	55	1556688.38	1508509.46	0.0319	508.73	15.05
p127	35	35	5	142755.25	140018.50	0.0195	47.23	1.53
p128	35	35	15	426582.36	424465.11	0.0050	123.36	3.82
p129	35	35	25	721405.25	696552.42	0.0357	249.37	6.96
p130	35	35	35	1000922.39	995382.32	0.0056	273.17	10.95
p131	35	35	45	1260729.03	1253779.84	0.0055	328.92	16.83
p132	35	35	55	1519463.22	1509069.08	0.0069	423.71	20.94
p133	35	45	5	153680.36	153244.71	0.0028	48.27	1.73
p134	35	45	15	451462.17	448229.10	0.0072	183.61	4.76
p135	35	45	25	733060.51	713025.63	0.0281	341.67	9.01
p136	35	45	35	1002440.40	992441.09	0.0101	341.88	14.25

p137	35	45	45	1370239.72	1361028.59	0.0068	437.70	18.84
p138	35	45	55	1665092.23	1640167.23	0.0152	580.55	22.28
p139	35	55	5	152254.38	151955.28	0.0020	63.37	2.02
p140	35	55	15	471759.92	453721.20	0.0398	250.73	5.89
p141	35	55	25	779455.85	762507.59	0.0222	385.10	10.32
p142	35	55	35	1136940.04	1108155.95	0.0260	604.52	12.48
p143	35	55	45	1370175.88	1360059.43	0.0074	530.78	24.06
p144	35	55	55	1657125.86	1641724.95	0.0094	657.01	36.06
p145	45	5	5	151464.83	150730.96	0.0049	10.34	0.82
p146	45	5	15	472442.90	468563.17	0.0083	35.92	1.71
p147	45	5	25	782592.84	775699.78	0.0089	50.54	2.03
p148	45	5	35	1132040.58	1122693.03	0.0083	78.11	2.78
p149	45	5	45	1479044.71	1469441.57	0.0065	104.17	3.81
p150	45	5	55	1745376.77	1662830.43	0.0496	270.69	4.58
p151	45	15	5	165597.96	163647.33	0.0119	30.51	1.13
p152	45	15	15	495683.83	475931.67	0.0415	131.32	2.42
p153	45	15	25	828454.27	822932.45	0.0067	115.66	4.19
p154	45	15	35	1049245.99	1040839.94	0.0081	168.29	6.67
p155	45	15	45	1510257.48	1496311.18	0.0093	216.38	9.52
p156	45	15	55	1782746.21	1771220.08	0.0065	252.62	13.80
p157	45	25	5	178792.69	178277.42	0.0029	40.07	1.51
p158	45	25	15	498868.43	496247.07	0.0053	106.60	3.85
p159	45	25	25	880172.03	841689.10	0.0457	367.95	7.08
p160	45	25	35	1174927.56	1165771.80	0.0079	252.59	11.29
p161	45	25	45	1509364.74	1497980.47	0.0076	325.33	13.17
p162	45	25	55	2001806.63	1917599.89	0.0439	771.55	19.73
p163	45	35	5	176519.16	175849.08	0.0038	51.76	1.74
p164	45	35	15	495853.54	492444.09	0.0069	146.60	4.77
p165	45	35	25	872130.30	866771.99	0.0062	259.33	9.07
p166	45	35	35	1191769.30	1161095.16	0.0264	496.72	13.54
p167	45	35	45	1513532.36	1491595.66	0.0147	510.09	18.23
p168	45	35	55	1978970.25	1966186.73	0.0065	540.51	29.41
p169	45	45	5	200551.19	196186.08	0.0222	108.05	2.25
p170	45	45	15	567274.17	556840.86	0.0187	237.33	6.12
p171	45	45	25	958868.15	951459.93	0.0078	306.17	10.55
p172	45	45	35	1231655.77	1223396.68	0.0068	414.84	17.71
p173	45	45	45	1600355.05	1590482.74	0.0062	558.40	18.39
p174	45	45	55	2064613.08	1998436.15	0.0331	957.82	29.49
p175	45	55	5	197118.02	196681.62	0.0022	84.22	2.84
p176	45	55	15	586436.69	566815.21	0.0346	420.30	6.83
p177	45	55	25	959689.29	954108.57	0.0058	404.69	11.70
p178	45	55	35	1303825.41	1272369.32	0.0247	705.73	20.67
p179	45	55	45	1673519.95	1661864.70	0.0070	667.45	30.81
p180	45	55	55	1989474.04	1971951.24	0.0089	886.21	45.16
p181	55	5	5	200613.67	199779.80	0.0042	19.51	1.05
p182	55	5	15	572199.02	565382.99	0.0121	42.54	1.71
p183	55	5	25	908802.98	903251.63	0.0061	61.75	2.32

p184	55	5	35	1398092.12	1389345.20	0.0063	86.68	3.57
p185	55	5	45	1788341.29	1781031.73	0.0041	113.25	4.41
p186	55	5	55	2229179.66	2211899.22	0.0078	162.33	6.36
p187	55	15	5	208679.86	207589.34	0.0053	37.05	1.48
p188	55	15	15	569567.49	565163.48	0.0078	95.24	3.18
p189	55	15	25	1105082.87	1053356.43	0.0491	1104.78	5.79
p190	55	15	35	1435059.80	1379879.29	0.0400	379.45	8.19
p191	55	15	45	1802186.02	1780600.72	0.0121	299.46	12.87
p192	55	15	55	2136237.64	2123560.75	0.0060	364.11	17.34
p193	55	25	5	207613.79	203183.37	0.0218	56.12	1.99
p194	55	25	15	610262.11	589526.84	0.0352	185.55	5.34
p195	55	25	25	1001128.43	995568.69	0.0056	263.09	7.94
p196	55	25	35	1434246.83	1421207.26	0.0092	371.01	12.97
p197	55	25	45	1920388.42	1839689.44	0.0439	870.10	14.76
p198	55	25	55	2266161.03	2249017.98	0.0076	473.67	19.96
p199	55	35	5	208258.50	208088.31	0.0008	64.68	2.41
p200	55	35	15	603928.99	599776.64	0.0069	182.69	5.15
p201	55	35	25	1071358.57	1053793.12	0.0167	368.15	11.33
p202	55	35	35	1514762.18	1474416.04	0.0274	559.20	18.51
p203	55	35	45	1921370.74	1909428.52	0.0063	531.84	25.27
p204	55	35	55	2489166.51	2372805.14	0.0490	1296.87	27.95
p205	55	45	5	213537.51	212834.37	0.0033	76.90	3.10
p206	55	45	15	680713.51	656160.75	0.0374	379.43	7.97
p207	55	45	25	1060249.27	1030749.79	0.0286	503.55	13.54
p208	55	45	35	1464018.98	1453328.94	0.0074	508.65	17.58
p209	55	45	45	2168615.30	2098559.01	0.0334	1304.45	26.64
p210	55	45	55	2378128.45	2368025.35	0.0043	832.32	44.82
p211	55	55	5	241093.34	233590.83	0.0321	156.14	3.22
p212	55	55	15	664143.87	660284.11	0.0058	279.78	6.51
p213	55	55	25	1114365.85	1087095.40	0.0251	528.42	16.50
p214	55	55	35	1554976.96	1542934.63	0.0078	652.62	22.69
p215	55	55	45	2026499.71	2001879.29	0.0123	1017.86	33.40
p216	55	55	55	2515776.05	2417971.23	0.0404	1782.84	38.57
————	—————	—————	Average	487340.42	480779.64	2.1416	222.75	7.39
————	—————	—————	Standard deviation	596521.20	587041.46	0.0332	267.89	8.48

Bibliography

Aissaoui, N., Haouari, M., and Hassini, E. (2007). Supplier selection and order lot sizing modeling: A review. *Computers & operations research*, **34**(12), 3516–3540.

Anisul Huq, F., Stevenson, M., and Zorzini, M. (2014). Social sustainability in developing country suppliers: An exploratory study in the ready made garments industry of bangladesh. *International Journal of Operations & Production Management*, **34**(5), 610–638.

Apple (2018). *Apple's Supplier Responsibility 2018 Progress Report*.

Azadnia, A. H., Saman, M. Z. M., and Wong, K. Y. (2015). Sustainable supplier selection and order lot-sizing: an integrated multi-objective decision-making process. *International Journal of Production Research*, **53**(2), 383–408.

Basnet, C. and Leung, J. M. (2005). Inventory lot-sizing with supplier selection. *Computers & Operations Research*, **32**(1), 1–14.

Bies, R. J., Bartunek, J. M., Fort, T. L., and Zald, M. N. (2007). Corporations as social change agents: Individual, interpersonal, institutional, and environmental dynamics. *Academy of Management Review*, **32**(3), 788–793.

Brammer, S., Jackson, G., and Matten, D. (2012). Corporate social responsibility and institutional theory: New perspectives on private governance. *Socio-economic review*, **10**(1), 3–28.

Breiman, L. (2001). Random forests. *Machine learning*, **45**(1), 5–32.

Brundtland, G. H. (1987). What is sustainable development. *Our common future*, pages 8–9.

Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., VanderPlas, J., Joly, A., Holt, B., and Varoquaux, G. (2013). API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122.

Carroll, A. B. (2008). A history of corporate social responsibility. In *The Oxford handbook of corporate social responsibility*.

Ciliberti, F., Pontrandolfo, P., and Scozzi, B. (2008). Investigating corporate social responsibility in supply chains: a sme perspective. *Journal of cleaner production*, **16**(15), 1579–1588.

Cook, S. A. (1971). The complexity of theorem-proving procedures. In *Proceedings of the third annual ACM symposium on Theory of computing*, pages 151–158. ACM.

Egan, M. L., Mauléon, F., Wolff, D., and Bendick, M. (2003). France's nouvelles regulations economiques: Using government mandates for corporate reporting to promote environmentally sustainable economic development. In *25th Annual Research Conference for Public Policy and Management, Washington, DC*, pages 6–8.

Florian, M. and Klein, M. (1971). Deterministic production planning with concave costs and capacity constraints. *Management Science*, **18**(1), 12–20.

Gates, D. (2013). Boeing 787's problems blamed on outsourcing, lack of oversight. *The Seattle Times Business/Technology*, **2**.

Governance & Accountability Institute, I. (2017). *Flash Report: 85% of S&P 500 Index Companies Publish Sustainability Reports in 2017*. [Online; accessed October-2018].

Hassini, E., Surti, C., and Searcy, C. (2012). A literature review and a case study of sustainable supply chains with a focus on metrics. *International Journal of Production Economics*, **140**(1), 69–82.

Ho, W., Xu, X., and Dey, P. K. (2010). Multi-criteria decision making approaches for supplier evaluation and selection: A literature review. *European Journal of operational research*, **202**(1), 16–24.

Kleindorfer, P. R., Singhal, K., and Wassenhove, L. N. (2005). Sustainable operations management. *Production and operations management*, **14**(4), 482–492.

McWilliams, A., Siegel, D. S., and Wright, P. M. (2006). Corporate social responsibility: Strategic implications. *Journal of management studies*, **43**(1), 1–18.

Militaru, G. and Ionescu, S. (2006). The competitive advantage of corporate social responsibility. *University" Politehnica" of Bucharest Scientific Bulletin, Series D: Mechanical Engineering*, **68**(2), 89–95.

Miner, G., Elder IV, J., and Hill, T. (2012). *Practical text mining and statistical analysis for non-structured text data applications*. Academic Press.

Rabieh, M., Babaee, L., Fadaei, R. A., and Esmaeili, M. (2018). Sustainable supplier selection and order allocation: An integrated delphi method, fuzzy topsis and multi-objective programming model. *Scientia Iranica*, pages –.

Schrijver, A. (1998). *Theory of linear and integer programming*. John Wiley & Sons.

Sills Egsgard LLP Article (nd̈). *Risks Arising from Violation of CSR Norms: Litigation, Social Licence and Reputational Damage*.

Tajbakhsh, A. and Hassini, E. (2015). Performance measurement of sustainable supply chains: a review and research questions. *International Journal of Productivity and Performance Management*, **64**(6), 744–783.

Tan, A.-H. *et al.* (1999). Text mining: The state of the art and the challenges. In *Proceedings of the PAKDD 1999 Workshop on Knowledge Disocovery from Advanced Databases*, volume 8, pages 65–70.

Toker, O. and Ozbay, H. (1995). On the np-hardness of solving bilinear matrix inequalities and simultaneous stabilization with static output feedback. In *American Control Conference, Proceedings of the 1995*, volume 4, pages 2525–2526. IEEE.

UPS CSR Report (2013). *UPS Corporate Sustainability Report 2013*. [Online; accessed October-2018].

Wagner, H. M. and Whitin, T. M. (1958). Dynamic version of the economic lot size model. *Management science*, **5**(1), 89–96.

WEF (2015). *Beyond Supply Chains Empowering Responsible Value Chains*. [Online; accessed October-2018].

White, G. (2017). What's changed since more than 1,110 people died in bangladesh's factory collapse. *URL: https://www. theatlantic. com/business/archive/2017/05/ranaplaza-four-years-later/525252/(3. 5. 2017).*

Widiarto Sutantoputra, A. (2009). Social disclosure rating system for assessing firms' csr reports. *Corporate Communications: An International Journal*, **14**(1), 34–48.

Zhang, H. (2004a). The optimality of naive bayes. *AA*, **1**(2), 3.

Zhang, T. (2004b). Solving large scale linear prediction problems using stochastic gradient descent algorithms. In *Proceedings of the twenty-first international conference on Machine learning*, page 116. ACM.