Integrating Data-Driven Forecasting and Large-Scale Optimization to Improve Humanitarian Response Planning and Preparedness

by

Tina Rezvanian

Submitted to the Department of Mechanical and Industrial Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Northeastern University September 2019
Abstract

This dissertation investigates the advantages of optimization and machine learning algorithms to characterize, predict, and solve Response Planning and Preparedness problems in large-scale humanitarian organizations. Organizations often base their operational decisions on the staff’s experiential knowledge rather than data-driven mechanisms. International professionals are one of the most valuable resources for humanitarian organizations that deliver food and relief items. The problem of assigning such personnel to positions based on their preferences is a two-sided stable matching problem. Many two-sided markets form a matching between their agents by running centralized clearinghouse algorithms that ensure “stable” and “perfect” assignments. When dealing with large-scale organizations, agents often are not aware of all of their options or inquiring about all candidates/positions can be costly. It is a well-known fact that a perfect matching in a system with partial information on agents’ preferences is incompatible with ensuring stability. To address this issue, we design cycle-based approximation mechanism that models negotiations between self-interested agents and identifies a greater-cardinality matching by relaxing stability minimally while maximizing social welfare for all agents across multiple matching cycles.

Similar to many other humanitarian organizations, UNHCR operates separate supply chains for the ongoing (OO) and emergency relief (ER) operations, which is costly and susceptible to variances in ER demand quantities. We design network configuration capable of matching humanitarian relief supplies with uncertain demands for the UNHCR’s supply chains. In doing so, a scenario-based two-stage stochastic modeling approach examines the possible actions for cost and lead time reduction such as merging OO and ER supply chains, moving global warehouses closer to demand locations, and incorporating influencing factors in warehouse selection decisions. We employ a multi-criteria decision making approach that makes inventory pre-position, shipping, and transportation mode decisions.

Shifting focus from reaction to anticipation of ER Response Planning and Preparedness, we propose a framework that improves the strategies for forecasting UNHCR’s ER response as a zero-inflated skewed variable. Working within the proposed framework, we explore the relationship between countries’ characteristics and the UNHCR emergency relief response. We predict countries’ future UNHCR ER response, the likelihood of different ER response levels, and the amount of ER response for a given country. In our analysis, identified patterns and associations determine the main drivers of ER response.
Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Two-stage Stable Matching with Negotiations</td>
<td>12</td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>12</td>
</tr>
<tr>
<td>1.2 Related Work</td>
<td>16</td>
</tr>
<tr>
<td>1.3 Case Study</td>
<td>19</td>
</tr>
<tr>
<td>1.3.1 Sets, Parameters, and Decision Variables</td>
<td>22</td>
</tr>
<tr>
<td>1.3.2 Mathematical Model</td>
<td>23</td>
</tr>
<tr>
<td>1.4 Problem Statement</td>
<td>25</td>
</tr>
<tr>
<td>1.5 Multi-period Matching Mechanisms</td>
<td>26</td>
</tr>
<tr>
<td>1.5.1 Repeating Stable Single-Period Assignments</td>
<td>26</td>
</tr>
<tr>
<td>1.5.2 Matching with Negotiations</td>
<td>28</td>
</tr>
<tr>
<td>1.5.3 Exact Solution for Perfect Relaxed-Stable Matching</td>
<td>30</td>
</tr>
<tr>
<td>1.5.4 Decision Variables</td>
<td>32</td>
</tr>
<tr>
<td>1.5.5 Mathematical Model</td>
<td>33</td>
</tr>
<tr>
<td>1.5.6 Cycle-Based Approximation Solution for Perfect Relaxed-Stable Matching</td>
<td>39</td>
</tr>
<tr>
<td>1.5.6.1 Pre-processing</td>
<td>39</td>
</tr>
<tr>
<td>1.5.6.2 Algorithmic method:</td>
<td>42</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>1.6</td>
<td>Computational Study</td>
</tr>
<tr>
<td>1.7</td>
<td>Conclusion</td>
</tr>
<tr>
<td>2</td>
<td>Integrating Supply Chains for Emergencies and Ongoing Operations in UNHCR</td>
</tr>
<tr>
<td>2.1</td>
<td>Introduction and Motivation</td>
</tr>
<tr>
<td>2.2</td>
<td>Literature review</td>
</tr>
<tr>
<td>2.3</td>
<td>Methodology for qualitative empirical study</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Case Selection</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Case Description</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Data Collection</td>
</tr>
<tr>
<td>2.3.4</td>
<td>Data Analysis</td>
</tr>
<tr>
<td>2.4</td>
<td>Model development based on empirical findings</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Merging Two Supply Chains</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Influencing factors identified in qualitative study</td>
</tr>
<tr>
<td>2.4.2.1</td>
<td>Demand Characteristics</td>
</tr>
<tr>
<td>2.4.2.2</td>
<td>Logistics</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Model</td>
</tr>
<tr>
<td>2.4.3.1</td>
<td>Problem definition</td>
</tr>
<tr>
<td>2.4.4</td>
<td>Minimum Expected Total Cost Model</td>
</tr>
<tr>
<td>2.4.4.1</td>
<td>Minimum Expected Lead Time Model</td>
</tr>
<tr>
<td>2.4.4.2</td>
<td>Budget Constrained Minimum Expected Lead Time Model</td>
</tr>
<tr>
<td>2.4.5</td>
<td>Case for model testing</td>
</tr>
<tr>
<td>2.5</td>
<td>Computational Results</td>
</tr>
<tr>
<td>2.6</td>
<td>Conclusions, Implications, and Future research</td>
</tr>
</tbody>
</table>
3 Forecasting Refugees Emergency Relief Response at United Nations

High Commissioner for Refugees (UNHCR)

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Introduction and Motivation</td>
<td>86</td>
</tr>
<tr>
<td>3.2</td>
<td>Literature Review and Our Contribution</td>
<td>89</td>
</tr>
<tr>
<td>3.3</td>
<td>Data</td>
<td>91</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Publicly Available Data</td>
<td>91</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Data Collected Through Collaboration with UNHCR</td>
<td>95</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Extracting Refugee-Specific Information</td>
<td>95</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Exploratory Analysis of ER Response and Explanatory Variables</td>
<td>98</td>
</tr>
<tr>
<td>3.3.5</td>
<td>Limitations</td>
<td>105</td>
</tr>
<tr>
<td>3.4</td>
<td>Modeling with an Empirical Strategy: Forecast the Amount of ER Response in Each Country</td>
<td>106</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Feature Selection Based on Correlation Analysis</td>
<td>109</td>
</tr>
<tr>
<td>3.4.1.1</td>
<td>Correlation Analysis of the Initial Feature Set</td>
<td>109</td>
</tr>
<tr>
<td>3.4.1.2</td>
<td>Correlation Analysis with Additional Factors</td>
<td>112</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Benchmark Models and Their Comparison</td>
<td>112</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Training Module: Two-part Learning Algorithm for Forecasting ER Response</td>
<td>115</td>
</tr>
<tr>
<td>3.4.3.1</td>
<td>First Learning Stage: ER Response Occurrence Model</td>
<td>118</td>
</tr>
<tr>
<td>3.4.3.2</td>
<td>First Learning Stage: Connection to Data</td>
<td>121</td>
</tr>
<tr>
<td>3.4.3.3</td>
<td>Second Learning Stage: ER Response Intensity Model</td>
<td>123</td>
</tr>
<tr>
<td>3.4.3.4</td>
<td>Second Learning Stage: Connection to Data</td>
<td>124</td>
</tr>
<tr>
<td>3.4.4</td>
<td>Prediction Module</td>
<td>127</td>
</tr>
</tbody>
</table>
3.4.4.1 Predict Occurrence of Positive ER Responses using C1 Occurrence Model (C3) ....................................................... 127

3.4.4.2 Predict the Amount of ER response Using C2 Regression Model (C4) .............................................................. 129

3.5 Additional Analysis: Forecast the Severity Levels of ER Response in Each Country ..................................................... 130

3.6 Insights, Recommendations, and Additional Exploration of Results ................................................................. 134

3.7 Conclusion ......................................................................................................................................................... 137

A ........................................ 140

B ........................................ 142
List of Figures

1-1 Comparison of Matching from the RS committee meeting and the Mathematical Model [1.5.5] ................................................. 24
1-2 Trajectory of number of unsatisfied agents with correlated preferences ................................................................. 47
1-3 Trajectory of number of unsatisfied agents with random preferences ................................................................. 48
2-1 Framework for network design and factors to include in model development ................................................................. 61
2-2 Warehouses in UNHCR’s network ................................................................................................................................. 76
2-3 Variability in ER demand over 2011 - 2013 per country. ................................................................................................. 77
2-4 Comparison of current versus extended network ........................................................................................................... 79
3-2 UNHCR ER response expenditure from 2011 to 2016 in refugees’ residence countries ............................................................ 99
3-3 Frequency distribution of UNHCR ER response expenditure from 2011 to 2016 .............................................................. 99
3-4 UNHCR ER response expenditure in refugees’ residence regions .................................................................................... 101
3-5 Refugee movements between geographical regions ............................................................................................................. 102
3-6 Pairwise association among features and ER response ...................................................................................................... 103
3-7 Ranked-based pairwise association among features and ER response .............................................................................. 104

7
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-8</td>
<td>Proposed framework for forecasting ER response</td>
<td>109</td>
</tr>
<tr>
<td>3-9</td>
<td>Predictions from the Benchmark Model A: Actual versus predicted ER</td>
<td>113</td>
</tr>
<tr>
<td>3-10</td>
<td>Predictions from the Benchmark Model B: Actual versus predicted ER</td>
<td>114</td>
</tr>
<tr>
<td>3-11</td>
<td>Model Performance of Benchmark Models: Comparison of actual versus</td>
<td>115</td>
</tr>
<tr>
<td></td>
<td>predicted plots for benchmark models A and B</td>
<td></td>
</tr>
<tr>
<td>3-12</td>
<td>Comparison of predicted LR probabilities and predicted SVM classes</td>
<td>120</td>
</tr>
<tr>
<td>3-13</td>
<td>Model Performance of Occurrence Model C1: Number of correct identi-</td>
<td>121</td>
</tr>
<tr>
<td></td>
<td>fications versus missclassification</td>
<td></td>
</tr>
<tr>
<td>3-14</td>
<td>Model Performance of LR versus SVM Model: Total number of miss-</td>
<td>122</td>
</tr>
<tr>
<td></td>
<td>classification under different feature combinations</td>
<td></td>
</tr>
<tr>
<td>3-15</td>
<td>Model Performance of SVM Model: Number of missed countries and</td>
<td>123</td>
</tr>
<tr>
<td></td>
<td>false alarms under different feature combinations</td>
<td></td>
</tr>
<tr>
<td>3-16</td>
<td>Predictions of Gamma Regression Model C2: Actual versus predicted</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td>ER response for countries where positive ER responses were actually</td>
<td></td>
</tr>
<tr>
<td></td>
<td>incurred in 2016</td>
<td></td>
</tr>
<tr>
<td>3-17</td>
<td>Effect of features on occurrence of a positive ER response from the</td>
<td>126</td>
</tr>
<tr>
<td></td>
<td>SVM model</td>
<td></td>
</tr>
<tr>
<td>3-18</td>
<td>Effect of features on the amount of ER response given that positive</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>ER responses were actually incurred using Gamma regression model C2</td>
<td></td>
</tr>
<tr>
<td>3-19</td>
<td>Predictions from the Proposed Two-Part Learning Algorithm: Actual</td>
<td>129</td>
</tr>
<tr>
<td></td>
<td>versus predicted ER response for all countries</td>
<td></td>
</tr>
<tr>
<td>3-20</td>
<td>Model Performance of MC LR model: Confusion matrix</td>
<td>134</td>
</tr>
<tr>
<td>3-21</td>
<td>Predicted probabilities from MC LR model</td>
<td>139</td>
</tr>
</tbody>
</table>
## List of Tables

1.1 Network scenarios for computational study ........................................ 45
1.2 Comparison of Deferred Acceptance algorithm (DA), Integer Programming (IP) model and Heuristic (H) for markets with random ordering of preferences in acceptable items ........................................ 49
1.3 Comparison of Deferred Acceptance algorithm (DA), Integer Programming (IP) model and Heuristic (H) for markets with 90% correlation in preference ordering of acceptable items ........................................ 50
2.1 Overview of data collection .................................................................. 60
2.4 Location-specific percentages for logistics cost .................................... 77
2.5 Network configurations under different margins within the budget constraints ........................................ 81
2.6 Distribution of OO and ER volumes satisfied per warehouse location .... 82
3.1 Feature descriptions and summary of data before normalization ......... 92
3.4 Performance metrics of benchmark models A and B ............................ 114
3.5 Actual versus predicted ER response and the amounts of past ER and OO expenditures in false alarm countries ........................................ 128
3.6 History of ER and OO response values for missed countries ............... 128
3.7 Performance Metrics: Comparison of proposed model versus benchmark models ........................................ 130

3.8 Ranges associated with classes of severity levels ........................................ 130

3.9 Comparison of actual versus predicted severity classes from the Two-Part Learning model and MC LR model in countries with miss-aligned predictions ........................................ 133
Chapter 1

Two-stage Stable Matching with Negotiations

1.1 Introduction

Organizations often base their staff assignment on individual’s experience that is built up over time rather than using reliable mechanisms. It is important that organizations maintain a transparent policy for their staff assignment procedure. Such policies guarantee that the staff and the hiring managers understand the procedure by which matchings are determined. However, since staff and managers are self-interested participants, it is inevitable to observe strategic actions made by both sides that may influence the match results. Therefore, a matching mechanism should be immune to such efforts. A mechanism in which stating true preferences is the optimal strategy for both sides is usually referred to as a “strategy-proof” mechanism. Most human resource departments, currently, do not entirely consider managers’ and staff’s preferences in their staff assignment. Also, adopted mechanisms are not
efficient, transparent or strategy-proof. The National Residents Matching Program (NRMP) was one of the first attempts for quantifying the impact of reliable matching mechanisms on satisfaction levels. It was celebrated with a Nobel Prize in 2012 due its preeminent societal impact.

In labor markets, the assignment and reassignment of personnel to jobs, collectively called as “agents”, is a large-scale problem faced by many organizations including the military, multi-national non-profits such as United Nations agencies, medical organizations, and others. Military assignment of officers to branches are among the sectors that have multiple matching cycles in a year. Successive dissatisfaction of officers with their match leads in high turnovers and market failure [54].

**Two-sided Markets.** Two-sided matching markets, first introduced by Gale and Shapely in 1962, refer to markets where one class of agents are to be matched with another [37]. Matching of students to colleges, personnel to jobs, and marriageable men to women are some of the examples. In two-sided matching markets, agents of both sides have a preference ranking over each other. The stable marriage problem is a two-sided matching problem where each agent can be matched to only one agent in the opposite side or can remain unmatched. Many two-sided markets form a matching between agents by running centralized clearinghouse algorithms that ensure *stable* matchings, where there are no pairs of agents preferring each other to what they are assigned to. Stability is a property that ensures fairness and prevents agents from circumventing their centralized clearing house assignment through pairwise negotiation with one another. The reason being, there are no incentives for pairs to break their current matchings and form an assignment between each other in a decentralized format. Therefore the resulting match is *incentive compatible*.

**Periodic markets.** Stable Matching Problem has been extensively studied under various forms and settings. The application of proposed mechanisms has pro-
foundly improved quality of life for people across the world [1]. The National Residency Matching program, kidney exchange, and school choice are some the well-known projects. Surprisingly, there has been very limited focus on markets with more than one period. The straightforward idea for stable matching of same agents over multiple periods is to use any stable matching algorithm and implement it for each period. In the section about incomplete lists we explain why such an idea leads to market failure.

**Incomplete lists.** When dealing with large-scale systems, agents often have incomplete information over their options. They may not be aware of all of their options or making applications for all positions can be costly. It is a well-known fact that if preference lists are incomplete, then a perfect matching may not always exist. Gale and Sotomayor in 1985 showed that without loss of generality, when preference lists are consistent (person \( i \) has job \( j \) in its preference list if and only if job \( j \) has person \( i \) in its preference list) every stable matching has the same size and includes exactly the same agents.

As a result, stability in its traditional meaning leads in same agents remaining unmatched over all periods which, in turn, decreases their satisfaction level. Agents who end up unmatched after every matching cycle leave the system. High turnovers of agents over multiple periods is a market design failure.

Yet in practice, organizations need to assign all their personnel to available job positions at the end of every matching cycle. Having a periodic two-sided market requires many practical considerations to be adjusted over time, for instance, the notion of fairness, centralized utility, and agents’ satisfaction need to be redefined for periodic markets.

**Inconsistent preferences.** Another practical consideration for XYZ and many large organizations is that most personnel may prefer jobs that are in greater cities or
have greater salary, however such jobs may not accept all people who list them in their preference list. Consistent preferences as explained previously assumes that pairs agree on acceptance of each other. Such markets, when modeled as networks, represent undirected networks. Any stable matching exactly matches the same agents resulting in equal match sizes. On the contrary, inconsistent preferences can yield unequal match sizes and different matched agents over different stable matchings. Such markets can be represented by directed networks. Consider the a market with three personnel and three jobs. If all personnel list job 1 and 2 only, and jobs list all three people with any ordering, then the a personnel-proposing DA match has a match size of two, while job-proposing DA may assign all agents.

\[ p_1 > p_2, p_1 > p_3 > p_2, p_3 > p_2 \]

The difference in match sizes is at its maximum when different sides are proposing in a DA algorithm. Although matched agents under different stable matching can be different, this difference often is zero or very small. Therefore allowing such preferences does not mitigate the importance of this study. In section — we explain why this not restricting problem setting actually comes in handy for periodic markets.

**Objective.** One-to-one stable matching of personnel to jobs in multiple periods was a problem initially identified in a collaboration with large-scale Humanitarian Organizations. We examine this problem first from an operations research perspective and then through algorithmic methods. In this study, we define relaxed-stability and negotiation as two interconnected notions to guarantee a minimum level of stability and to increase match cardinality. These properties yield a trade-off point, on
which the stability of a period can be relaxed for the sake of perfectness.

The first section includes an integer programming model, maximizing total utility gained by all agents through all periods. The problem is constrained to flow balance, relaxed-stability and negotiation constraints. Note that the utility is a ranked based function. The second section provides a heuristic that identifies nodes that are the subject of negotiations and analyzes the aftermath accepting negotiations through cycles defined on both periods.

1.2 Related Work

Research originating from operations research has developed multi-period assignments to solve job rotation problems [12] and scheduling medical residents [36] using optimization and design of heuristics. However, these models are not defined for a two-sided setting with self-interest entities and do not generally include the notion of stability.

**Relaxation of Stability.** In the field of algorithmic game theory, relaxation of stability in order to attain larger match size is a practice that has been previously studied for single period markets [14][41]. Several studies has examined single period matching markets when stability is relaxed only to attain other properties as well such as keeping couples together [59][51] and incorporating utility functions [4].

Papers interested in increasing match size are more related to the purpose of our study. Since it is not possible to maintain stability while forming a perfect matching, one has to be relaxed for the sake of the other. Relaxation of stability in order to attain larger match size is a practice that has been previously studied for single period markets. [11] presented a hospital/resident matching problem of 781 students to be matched to 789 positions in 53 hospitals. The Scottish Foundation Allocation
Scheme (SFAS) found a stable matching with 37 unmatched, where administrators offered relaxation of stability in order to attain a larger matching. They found a matching that has maximum cardinality while it is “as stable as possible” with 10 blocking pairs [14].

**Dynamic Markets.** Several researchers have studied stable matching over multiple periods with dynamic markets over the recent years. Research on dynamic markets mainly focuses on markets with two periods, where agents’ preferences can change over different periods. In a two-period market, preference lists strictly order pairs of agents rather than individual agents. This corresponds to the possibility of a person preferring job $j$ to job $k$ in the first period and vice versa second period. Damiano and Lam in 2005 and Kurino in 2009 analyze alternative forms of stability adjusted to dynamic markets [24, 53]. Also, Kadam in 2014 and Kotowski in 2015 both refer to strict stability over all periods as “dynamically stable” and “robust prescient stable” matchings and show that with some conditions on the structure of the preference lists, there always exist a dynamically stable matching that can be identified by modified Deferred Acceptance algorithm [50, 52]. Although dynamic markets embed multiple periods, it is not the focus of our study as they adopt a different format for ordering preferences and allow agents to change their preferences over agents of the opposite side.

**Dependent Matching.** Sönmez and Switzer in 2013 analyze stable matching of cadets to branches where cadets are allowed to bid on their favorite branches through terms of commitments. Matchings in this setting are formed through contracts with multiple terms. Stability is defined such that 1) for any cadet, branch and term of commitment selected in a contract, there should not exist a cadet or a branch that prefer to reject the contract, 2) there should not exit a cadet, branch and term of commitment not selected in a contract where the cadet prefer that branch and term
of commitment but also the contract has sufficiently high priority to be selected by the branch.

Sönmez and Switzer in 2013 analyze stable matching of cadets to branches where cadets are allowed to bid on their favorite branches through terms of commitments. Matchings in this setting are formed through contracts with multiple terms. Stability is defined such that 1) for any cadet, branch and term of commitment selected in a contract, there should not exist a cadet or a branch that prefer to reject the contract, 2) there should not exit a cadet, branch and term of commitment not selected in a contract where the cadet prefer that branch and term of commitment but also the contract has sufficiently high priority to be selected by the branch.

Note that, in Sönmez and Switzer work, any term of commitment is dependent to the branch in a contract. In our work, the assignment of each agent at each period also depends on its assignment in the other period. Therefore, we can look at a contract with a branch and only one term as a matching over two periods where the branch is the matching of cadet for the first period, and the term of commitment is the matching of cadet for the second period. Then we can define our work as an extension of this paper only with terms of contracts identified by a different strategy, called negotiations, that pursuit resolution of issues with stable matching in large periodic markets allowing incomplete preference lists. Therefore, we place our study at the interface of 1) [14], 2) Sönmez and Switzer and 3) dynamic markets.

**Complexity and Polyhedral Analysis.** In 1989 initially proved that the stable matching problem has an integral polyhedron. He also showed that the stable matching polytope is the convex hull of the stable marriage solutions. As a result, stable matching can be found through linear programming. Their note in 1962 showed that when the lists of preferences order every agent of the other side (complete preference lists), there always exists at least one stable matching
that is perfect and can be found in $O(n^2)$ time for an instance involving $n$ agents in a bipartite graph [37]. Deferred Acceptance algorithm (DA) is a certificate that the traditional stable matching is in the class of P [2, 76]. Later in 1986, Irving and Leather showed that the number of stable matchings when preference list are complete can be exponential in the size of problem [45].

Although still polynomial, the complexity changes significantly when preferences have different structures and incomplete preference list are allowed [40]. Note that the simplest and most complex scenarios occur when there is a wide divergence of opinion amongst the proposers. Holding identical opinions is intermediate in complexity [5]. [14] show that finding a maximum cardinality minimum blocking pair one-to-one stable matching with incomplete lists in a single-period setting is NP-hard and not approximable within $n^{1-\varepsilon}$ for any $\varepsilon > 0$.

The problem we are studying does not have an integral polytope since linear relaxation does not yield integral solution, and is NP-hard as it can be reduced to the 3-coloring problem. Also, our problem is a multi-period representation of finding a maximum cardinality minimum blocking pair problem, then our problem which restricts the number of blocking pairs through negotiations is also NP-hard.

1.3 Case Study

We partnered with an anonymous Humanitarian Organization, henceforth referred to as HO XYZ, that delivers humanitarian aid and relief to refugees across the world. HO XYZ, which is one of the largest agency in the world, has over 14,000 personnel distributed across headquarters and 80 other countries. A subset of over 1500 international professional staff members rotate positions every 2-4 years at an estimated reassignment cost of approximately $40,000 per person. This organization experi-
ences large turnovers among staff as a large subset of personnel end up unsatisfied with their assigned position at end of every matching cycle. As a results, any system that promotes fairness and welfare over consecutive assignment cycles are expected to increase satisfaction and reduce turnover.

HO XYZ runs 2-4 matching cycles per year finding assignments between approximately 500 personnel and positions. XYZ goal is to assign all personnel and positions. In contrast to the classical two-sided markets, in HO XYZ, there are staffing coordinators who take the role of a mediator between managers and Personnel. Personnel and managers unofficially negotiate with staffing coordinators to get matched to their desired option. Personnel, managers, and staffing coordinators indicate preference over their at most 8, 3, 5 agents, respectively. Staffing coordinators are aware of the position requirements and personnel background. Staffing coordinators “link” some of the positions to personnel based on negotiations before the matching process. Though the matching of such pairs is not mandatory, their assignment is always considered. In addition, there are certain positions that are in priority and need to be matched to one of their preferred candidates. Such pairs (staff and position) are refereed to as “priority links”. The matching process takes approximately 6 month and uses many resources. Currently, a large group of personnel and positions remain unsatisfied with their assignment (unmatched or matched to an agent not in their preference list). As a result, they either leave the system or re-enter the next matching cycle before their term of service is fulfilled.

Before matching personnel to positions, the preferences of managers and staffing coordinators need to be merged so that a two-sided matching mechanism is implemented. We use a weighed sum formulation, where weights are adjustable. In out analysis, staffing coordinators’ preference has a weight of 0.6 and managers 0.4. Having incomplete preferences with no minimum quota leads to challenges such as
sparsity of preferences. In this case, managers or personnel who did not list any candidate in their preference list remains unmatched in an stable matching.

The data of our case study includes 521 positions and 593 personnel who entered the first matching cycle of the in 2019. We adjusted the data per our discussions with our partners to merge the preferences of staffing coordinators and managers. This merged list represents position’s preference list. In merging Some of the adjustments are as follows:

- If a position is linked to a professional, but the staffing coordinator has not ranked him/her, then the linked professional is attached to the end of staffing coordinator’s preference list.

- If a position is linked to a professional and there are no positions listed at all by the professional, then the linked position is ranked as the 1st option.

- If a position is linked to a professional and but the personnel has not ranked the linked position (but has other positions listed), then, the linked position is attached to the end of professional’s list, unless the professional has declared he/she is not in agreement.

Currently, HO XYZ matches their personnel to the available positions in meetings, where the reassignment (RS) committee make the final decisions based on their discussions and comments submitted by the staffing coordinators. These meetings can take more than six months depending on the number of professional and positions. One shortcoming of the current system is that, the majority of personnel are matched to a position that is not listed in their preference list. As the number of listed options or candidates in preference list are shorter, the matching process takes more time and the in the aftermath comes additional negotiations.
Preference lists are incomplete and extremely short in comparison to the total number of personnel and positions. Also, manning constraints (priority links, rejected or added pairs) can violate the stability constraints. We develop a mathematical model that maximizes the utility gained by all agents and takes into account the manning constraints such as priorities. We create a tool that allows our partners to accept or reject a match once they are in the meetings. The goal is to use the model as a tool to see the best matchings at each stage of the committee meeting. In this mathematical model, the stability constraint are relaxed only if there is a manning constraints or an order is sent in the tool by the user. However, each time the stability is relaxed, a penalty is added to the objective function leading in a decline in the total utility gained by the matching.

1.3.1 Sets, Parameters, and Decision Variables

Sets
\[ S : \text{ set of personnel} \]
\[ P : \text{ set of positions} \]
\[ E : \text{ set of edges } (i,j) \text{ where personnel } i \text{ and position } j \text{ rank each other in their preference lists} \]
\[ A : \text{ set of edges } (i,j) \text{ that are accepted by a user} \]
\[ U : \text{ set of edges } (i,j) \text{ that are rejected by a user} \]

Parameters
\[ \pi^i_j : \text{ rank of position } j \text{ in personnel } i \text{’s list} \]
\[ \pi^j_i : \text{ rank of professional } i \text{ in position } j \text{’s list} \]
\[ q_s : \text{ maximum rank of positions in personnel list} \]
\( q_p \): maximum number of personnel in positions list \( j \)

\( d_{ij} \): 1, if the \((i,j)\) is a priority link; 0, otherwise

**Decision Variables**

\( x_{ij} \): 1, if professional \( i \) is matched to position \( j \), 0, otherwise

\( \delta_j \): integer \( \in \{-1, 0\} \)

### 1.3.2 Mathematical Model

\[
\text{Max} \quad \sum_{i \in S, j \in P} \left( \frac{q_p - \pi^i_j}{q_p} \right) x_{ij} + \sum_{i \in S, j \in J} \left( \frac{q_s - \pi^i_j}{q_s} \right) x_{ij} + \sum_{j \in P} \delta_j \quad (1.2a)
\]

Subject to

\[
\sum_{j \in J} x_{ij} \leq 1, \forall i \in S, \quad (1.2b)
\]

\[
\sum_{i \in S} x_{ij} \leq 1, \forall j \in P, \quad (1.2c)
\]

\[
\sum_{i \in S} x_{ij} \cdot d_{ij} - 1 = \delta_j, \forall j \in P, \quad (1.2d)
\]

\[
x_{ij} = 1, \forall (i, j) \in A, \quad (1.2e)
\]

\[
x_{ij} \neq 1, \forall (i, j) \in U, \quad (1.2f)
\]

\[
x_{ij} + \sum_{(i,k) \in E: \pi^i_k \leq \pi^i_j} x_{ik} + \sum_{(g,j) \in E: \pi^j_g \leq \pi^i_j} x_{gj} \geq 1, \forall (i, j) \in E \setminus (A \cup U), \quad (1.2g)
\]

In our case study, the number of positions in this matching cycle is less than the number of personnel. Consequently, if preference lists where complete, the maximum number of matching would be 521. In this data, only 360 of total positions are actually listed by any candidates. The remaining 161 positions are then not considered in
Similarly, out of the total 593 candidate personnel, only 415 have been listed as a candidate for a position by a manager of a staffing coordinator. Therefore, developed model excludes the remaining 178 from its matching decisions.

Feeding the preference lists to the developed model and comparing the resulting match with the decisions made by the RS committee at HO XYZ, Table 1-1 shows the relative performance of the two mechanisms. Although, RS committee provides a larger match size (400 pairs), 140 of the assignments are made without considering the preferences. such assignments are considered unacceptable by the candidates or positions side. The mathematical model, however, makes 337 assignments, all of which are acceptable by both sides. The mathematical model leaving 23 of the positions ,which at least provided one candidate in their preference list, unmatched. The RS committee’s matching leaves 48 of such positions unmatched. Therefore, although the total number of matchings from the RS committee is larger then the model, assignments found by Model 1.3.2 has a significantly higher utility.

Due to the large number of unmatched agents (personnel and positions) and unacceptable assignments, HO XYZ has investigated different approaches for increasing the total utility and satisfaction. On of such approaches has been aftermarket matching cycles, where more of the negotiations between personnel and the staffing coordinators takes place in. In other cases, members who are not satisfied with their assignments and have not fully served their time enter the aftermarket matching

<table>
<thead>
<tr>
<th>Matching Mechanism</th>
<th>Number of Positions</th>
<th>Number of Staff</th>
<th>Number of Matched Pairs</th>
<th>Number of Unmatched Positions</th>
<th>Number of Unacceptable Pairs Matched</th>
</tr>
</thead>
<tbody>
<tr>
<td>RS Committee</td>
<td>521.00</td>
<td>593.00</td>
<td>400.00</td>
<td>121.00</td>
<td>140.00</td>
</tr>
<tr>
<td>Model 1.3.2</td>
<td>521.00</td>
<td>593.00</td>
<td>337.00</td>
<td>184.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Figure 1-1: Comparison of Matching from the RS committee meeting and the Mathematical Model 1.5.5
cycles as well. However, using such aftermarket strategy separately from the main matching process would be mainly consisted of the positions and candidates that did not list each other in their preference list in the first matching cycle. Even with the additional members from previous cycles, the matching remains as a secondary-level assignment. Therefore, unless agents change their preferences, the same candidates and positions remain unsatisfied. It is needless to say that such a mechanism incentivize strategizing. In the next sections of this work, we investigate other approaches to identify a perfect matching between personnel and positions without leaving a large group unsatisfied.

### 1.4 Problem Statement

Let $P$ and $J$ denote finite and disjoint sets of personnel $P = \{p_1, p_n\}$ and jobs $J = \{j_1, j_n\}$. Each person (job) strictly orders its preferences over a subset of jobs (people) with unrestricted size which we refer to as acceptable lists. Acceptable lists are denoted by $\succ_i$ for all agents $i \in P \cup J$. The remaining subset of agents on the opposite side are strictly ordered in another set called unacceptable list denoted by $\not\succ_i$ for all agents $i \in P \cup J$. If $k$ is the number of matching cycles in of multi-period market, then let $k$ be equal to two through out mechanism design for the sake of simplicity.

In the classical one-to-one stable matching problem, a matching $\mu$ is a function that assigns each person to a job or leaves it unmatched $\mu : P \cup J \to J \cup P \cup \emptyset$ such that it is individually rational and stable. A matching $\mu$ is *individually rational* if there is no person or job that prefer remaining unmatched to its assignment under $\mu$. A matching $\mu$ is *stable* if there is no pair preferring each other to their match under $\mu$. The people-proposing deferred acceptance algorithm finds a stable matching through
a set of proposals from personnel to their most favorite jobs at each iteration. Each job rejects those not listed in its acceptable list and keeps only one job that is ranked best. The algorithm terminates when every agent has a match or once the acceptable lists are exhausted. The solution of deferred acceptance algorithm favors agents of the proposing side. In particular every agent of the proposing side are weakly better off in their own side proposing DA stable matching in comparison to any other stable matching [37]. Throughout the paper, we use the term “own-side proposing DA algorithm” own-side proposing DA algorithm” for when the proposing side is the side of the agent under consideration. If the proposing side for DA algorithm is not specified, then the statement holds regardless.

This study intends to extend the one-to-one stable matching problem to a multi-period setting. Based on our experience it is necessary to look at problems with inconsistent and incomplete preferences. This project is identified as real-world problem in the World Food Programme. Having inconsistent preferences means if person $p$ lists job $j$ in its acceptable list, job $j$ may not have person $p$ in its acceptable list. Inconsistent preferences can result in unequal match sizes and slightly different agents that are matched under different stable matchings.

### 1.5 Multi-period Matching Mechanisms

#### 1.5.1 Repeating Stable Single-Period Assignments

One mechanism is to use stable matchings for each period. Inconsistent preferences can result in slightly different match sizes across stable matchings. Consider a market with 3 personnel and 3 jobs and following preferences.
Personnel-proposing stable matching $\mu^P$ has a match size of 2 while the job-proposing stable matching has a match size of 3.

$$\mu^P = \left( p_1, p_2, p_3 \right) \quad \text{and} \quad \mu^J = \left( p_1, p_2, p_3 \right)$$

A greater match size corresponds to smaller unsatisfied agents. Let $\mu^{DA}$ to be equal to the DA stable matching whose size is greater.

$$\mu^{DA} \leftarrow \arg\max_{\mu^P, \mu^J} |\mu|$$

Equating the matching of all periods $\mu^{DA,k}$ to be equal to $\mu^{DA}$ results in minimum unsatisfied agents under a stable setting.

$$\text{if } \left( \mu^{DA,k=1} = \mu^{DA} \text{ and } \mu^{DA,k=2} = \mu^{DA} \right) \Rightarrow \left( \mu^{DA} \leftarrow \mu^{DA,1} \cup \mu^{DA,2} \right)$$

This precisely means that the matching of all $k$ periods is equal under DA algorithm. Note that, the difference in match size and matched agents not only depends how much agents of opposite side agree on the level of attractiveness for each other but also depends on the length of acceptable lists and correlation in each side’s preferences.
In large markets and short preference lists, using the solution to different stable matching algorithms at each period, leads in the majority of same agents remaining unmatched over all periods. Assuming there are two periods in the market, we are trying to find a point on which the stability can be relaxed in order to increase match cardinality. More precisely, the goal is to increase match cardinality over two periods such that the number of unmatched agents in both matching cycles is reduced.

1.5.2 Matching with Negotiations

Before explaining our proposed strategy for identifying a threshold for relaxing stability for the sake of match cardinality, we define some concepts and terminologies to adjust stability for matching markets with multiple periods.

Let $\mu_{H,k}$ be the match from our proposed heuristic at period $k$ and Let $\mu_{H,k}$ be the final matching of proposed heuristic across $k$ periods. Let $F_i$ denote the frequency of agent $i$ appearing in the opposite side’s acceptable list. We use the term “unsatisfied agents” for those who are either unmatched or matched to an item not in their acceptable list. We also interchangeably use terms agents and nodes as agents in a market correspond to nodes in a network.

**Definition 1.** If there is an agent $i$, unmatched under DA algorithm $\mu^{DA}$, whose frequency is greater than assignment of any of its acceptable items under DA, then we call node $i$ a deserving node.

\[ \forall i, j \in (\succ_i) \ s.t. \ \mu^{DA}(i) = i, \mu^{DA}(j) \neq j, \ if \ F_i > F_{\mu^{DA}(j)} \Rightarrow i \in DN \]

**Definition 2.** A matching $\nu : \nu^{k=1} \cup \nu^{k=2}$ is relaxed-stable over two periods if there does not exist a pair in the DA matching $\mu^{DA}$ that strictly prefer each other to both of their assignments under $\nu$. 

28
Definition 3. If $j$ is matched is DA matching and $i$ is a deserving node who is matched to $j$ under $\mu^{H,k}$, then period $k$ is agent $j$’s price period as it strictly prefers its DA matching to deserving node $i$.

$$\forall i \in DN, j \in (\succ_i) \text{ s.t. } \mu^D(j) \neq j, \text{ if } \mu^{H,k}(i) = j \Rightarrow k \text{ is Price Period for } j$$

The intuition is that agent $j$ is paying a price by accepting to be matched to deserving node $i$ which he prefers less than its DA match.

Definition 4. If period $k$ is the price period for agent $j$, then period $K'$ is agent $j$’s prize period, where it must receive a reward which is at least as preferred as its own-side proposing DA match.

Definition 5. Negotiation is a directed path over two periods that matches deserving agent $i$ to agent $j$ in its acceptable list at period $k$ only if there is a reward for agent $j$ in period $k'$.

The idea is increase match size over two periods only if accepting a negotiation increases total welfare of the market. A negotiation contract is offered to an agent who is desired by a deserving node. Since the negotiation is offered to a node who pays a price by being matched to a deserving agent at one period, then it must receive a reward in the other period for a negotiation to be completed.

Along side increasing match cardinality, the goal of negotiations is to reduce the number of agents that are unsatisfied over two periods in a socially acceptable and systematic way.
1.5.3 Exact Solution for Perfect Relaxed-Stable Matching

We propose a mathematical model for assignment and re-assignment of staff to jobs with incomplete and inconsistent preference lists. You can find the mathematical model in appendix of this chapter. In this model, each person \((i \in P)\) or job \((j \in J)\) is matched through one of the following match types at any period: 1) both-sided acceptable \((j \in (\succ p), p \in (\succ j))\), 2) one-sided acceptable \((j \in (\succ p), p \notin (\succ j))\) or 3) unacceptable \((j \notin (\succ p), p \notin (\succ j))\). In addition to variables associated to each match type, we define auxiliary binary variables to keep track of the sequence of price and prize periods when a negotiation is admitted.

**Objective function.** Developed mathematical model uses a ranked based function to calculate total welfare across two periods. In Model \([1.5.5]\) the objective \((F1)\) maximizes the total utility gained by all agents over two periods with different match types. If an agent is matched to an unacceptable pair, then utility gained is negative. Each time a negotiation is accepted, a benefit term is added to total utility. This benefit term is equal to utility of the match type plus to the frequency of the deserving node in negotiation.

**Constraints.** The set of constraints \((C1)\) to \((C4)\) in Model \([1.5.5]\) ensure exactly one assignment for each agent at any period. Constraints \((C5)\) to \((C14)\) allow relaxation of stability on pair \((k, s)\) in one period only if either one of the following is true: 1) for any pairs \((g, s)\) such that \(g\) is in \(k\)’s side, has \(s\) in its acceptable list and has greater frequency than \(s\) or 2) for any pair \((k, g)\) such that \(g\) is in \(s\)’s side, has \(k\) in its acceptable list and has greater frequency than \(k\). We call the set of constraints that ensure this property as relaxed-stability constraints. Note that in the classical one-to-one stable matching problem, the integer programming model defines stability constraints only for all both-sided acceptable pairs. That is because
preference lists in that problem setting are complete and the matching market can be represented by an undirected graph. On the contrary, the market we are studying also embeds one-sided acceptable pairs and represents a directed graph. Therefore, relaxed-stability constraints are defined for all one-sided and both-sided acceptable pairs.

The set of active relaxed-stability constraints for pair \((k, s)\) depend on whether agents in that pair are involved in a negotiation and if they are, what is the sequence of price and prize periods for the agent with greater frequency than either \(k\) or \(s\). Auxiliary binary variables keep track of admitted negotiations involving each pair in the market.

The next set of constraints enable negotiations along with relaxed-stability constraints. Constraints \((C16)\) to \((C22)\) enable negotiations along with relaxed-stability constraints. Negotiation constraints only allow a reward to be gained by an agent in one period if it is in its price period in the other period. On the other hand, relaxation of stability in one period opens that chance for a deserving node to be matched only if there is a reward for it in the other period. We take advantage of additional constraints \((C23)\) to \((C42)\) to track the sequence of price and prize periods for each negotiation.

Developed mathematical model finds the optimal solution for the problem of finding a relaxed-stable matching while maximizing the total social welfare. Developed IP model is never going to be infeasible due to including strictly ordered unacceptable lists and associated match types. However, since the problem is NP-hard, we cannot solve the IP model with large problem sizes. It is clear that the set of feasible solutions from IP model is large and the required memory and computational power for exhausting feasible solutions increases by problem size. The largest problem that we were able to solve with CPLEX was when \(n=35\). We intend to design a mech-
anism that does not look at all feasible solutions but systematically finds a feasible matching for the IP model that yields close-to-optimal total utility value.

### 1.5.4 Decision Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Constraint</th>
<th>Constraint</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_{ij} )</td>
<td>( i \in \succ j, j \in \succ_i )</td>
<td>( z_{ij} )</td>
</tr>
<tr>
<td>( x_{ij}^J )</td>
<td>( i \in \succ j, j \in \succ_i )</td>
<td>( z_{ij}^J )</td>
</tr>
<tr>
<td>( x_{ij}^W )</td>
<td>( i \in \succ j, j \in \succ_i )</td>
<td>( z_{ij}^W )</td>
</tr>
<tr>
<td>( f_{ij} )</td>
<td>( i \in \succ j, j \not\in \succ_i )</td>
<td>( h_{ij}^J )</td>
</tr>
<tr>
<td>( f_{ij}^W )</td>
<td>( i \not\in \succ j, j \in \succ_i )</td>
<td>( h_{ij}^W )</td>
</tr>
<tr>
<td>( f_{ij}^N )</td>
<td>( i \not\in \succ j, j \not\in \succ_i )</td>
<td>( h_{ij}^N )</td>
</tr>
<tr>
<td>( y_{ij} )</td>
<td>( i \in \succ j, j \not\in \succ_i )</td>
<td>( t_{ij} )</td>
</tr>
<tr>
<td>( y_{ij}^W )</td>
<td>( i \not\in \succ j, j \in \succ_i )</td>
<td>( t_{ij}^W )</td>
</tr>
<tr>
<td>( y_{ij}^J )</td>
<td>( i \in \succ j, j \not\in \succ_i )</td>
<td>( t_{ij}^J )</td>
</tr>
<tr>
<td>( y_{ij}^W )</td>
<td>( i \not\in \succ j, j \in \succ_i )</td>
<td>( t_{ij}^W )</td>
</tr>
</tbody>
</table>
1.5.5 Mathematical Model

\[
\begin{align*}
\max & \sum_{i \in J, j \in R_i} (v_{ij} + u_{ij})x_{ij} + \sum_{i \in J, j \in R_i} (F_i + v_{ij} + u_{ij})x_{ij}^W + \\
& \sum_{i \in J, j \in R_i} (F_j + v_{ij} + u_{ij})x_{ij}^J + \sum_{j \in R_i, i \in P_j} (u_{ij} + F_j + v_{ij})y_{ij}^J + \\
& \sum_{j \in R_i, i \in P_j} (v_{ij} + F_i + v_{ij})y_{ij}^W + \sum_{k \in R_i, i \in P_k} (v_{ij} + u_{ij})f_{ik}^W + \\
& \sum_{k \notin R_i, i \in P_k} (v_{ij} + u_{ij})f_{kj}^J + \sum_{i \in J, j \in R_i} (v_{ij} + u_{ij})z_{ij}^J + \\
& \sum_{i \in J, j \in R_i} (v_{ij} + F_i + v_{ij})z_{ij}^W + \sum_{i \in J, j \in R_i} (v_{ij} + F_j + v_{ij})z_{ij}^J + \\
& \sum_{k \in R_i, i \in P_k} (v_{ij} + F_i + v_{ij})t_{ik}^W + \sum_{k \in R_i, i \in P_k} (v_{ij} + F_j + v_{ij})t_{kj}^J + \\
& \sum_{k \notin R_i, i \in P_k} (v_{ij} + u_{ij})h_{ik}^W + \sum_{k \notin R_i, i \in P_k} (v_{ij} + u_{ij})h_{kj}^J + \\
& \sum_{k \notin R_i, i \in P_k} (v_{ij} + u_{ij})h_{ik}^N + \sum_{j \notin R_i, i \in P_j} (v_{ij} + u_{ij})f_{ij}^N \\
\end{align*}
\]

Subject to:

Every node must have exactly one match

\[
\begin{align*}
\sum_{i \in J, j \in R_i} x_{ij}^{MM} + \sum_{j \notin R_i, i \in P_j} y_{ij}^J + \sum_{j \notin R_i, i \in P_j} y_{ij}^W + \sum_{j \notin R_i, i \in P_j} y_{ij}^N = 1, \forall j \in J, \quad (1.4a) \\
\sum_{i \in J, j \in R_i} x_{ij}^{MM} + \sum_{j \notin R_i, i \in P_j} y_{ij}^J + \sum_{j \notin R_i, i \in P_j} y_{ij}^W + \sum_{j \notin R_i, i \in P_j} y_{ij}^N = 1, \forall i \in P, \quad (1.4b) \\
\sum_{i \in J, j \in R_i} z_{ij}^{MM} + \sum_{k \notin R_i, i \in P_k} t_{ik}^W + \sum_{k \notin R_i, i \in P_k} t_{kj}^J + \sum_{j \notin P \cup D, i \in P_j} t_{ik}^N = 1, \forall j \in J, \quad (1.4c) \\
\sum_{i \in J, j \in R_i} z_{ij}^{MM} + \sum_{k \notin R_i, i \in P_k} t_{ik}^W + \sum_{k \notin R_i, i \in P_k} t_{kj}^J + \sum_{j \notin P \cup D, i \in P_j} t_{ik}^N = 1, \forall j \in J, \quad (1.4d)
\end{align*}
\]

33
Subject to:

**Relaxed-stability constraints:**

**Forwards Negotiations**

**Pay period:**

\[
x_{ks} + \sum_{k \in \gamma, g \in \gamma_k} x_{kg} + \sum_{g \in \gamma_k, i \in \gamma_g} x_{gs} + \sum_{g \in \gamma_k, j \in \gamma_g} y^j_{gs} + \sum_{g \in \gamma_k, k < k_g} y^W_{kg} + \sum_{g \in \gamma_k, k < k_g, j \in \gamma_g} f^j_{gs} + \sum_{g \in \gamma_k, k < k_g, g \in \gamma_k} f^W_{kg} \geq 1.0_{ks}, \forall k \in \gamma_k, s \in \gamma_{ks}, \quad (1.5a)
\]

**Reward period:**

\[
z_{ks} + \sum_{i \in \gamma_k, j \in \gamma_i} z^M_{is} + \sum_{i \in \gamma_k, k \in \gamma_k} z^M_{k\gamma} + \sum_{j \in \gamma_k, k \in \gamma_k} z^M_{k\gamma} + \sum_{i \in \gamma_k, g \in \gamma_k} z^M_{k\gamma} + \sum_{g \in \gamma_k, k \in \gamma_k} z^M_{k\gamma} + \sum_{g \in \gamma_k, k \in \gamma_k} z^M_{k\gamma} + \sum_{g \in \gamma_k, k \in \gamma_k} z^M_{k\gamma} + \sum_{g \in \gamma_k, k \in \gamma_k} z^M_{k\gamma} \geq 1.0_{ks}, \forall k \in \gamma_k, s \in \gamma_{ks}, \quad (1.5b)
\]

**Backwards Negotiations**

**Reward period:**

\[
x_{ks} + \sum_{i \in \gamma_k, j \in \gamma_k} x^M_{is} + \sum_{i \in \gamma_k, j \in \gamma_k} x^M_{kj} + \sum_{j \in \gamma_k, g \in \gamma_k} x^M_{jk} + \sum_{g \in \gamma_k, k \in \gamma_k} x^M_{kg} + \sum_{g \in \gamma_k, k \in \gamma_k} x^M_{kg} + \sum_{g \in \gamma_k, k \in \gamma_k} x^M_{kg} + \sum_{g \in \gamma_k, k \in \gamma_k} x^M_{kg} \geq 1.0_{ks}, \forall k \in \gamma_k, s \in \gamma_{ks},
\]
\[ \sum_{g \in \pi, s \in \pi_g} y_{gs}^j + \sum_{g \in \pi, k \in \pi_g} y_{kg}^W + \sum_{g \in \pi, s \in \pi_g} f_{gs}^j + \sum_{g \in \pi, k \in \pi_g} f_{kg}^W \geq 1. v_{ks}, \quad \forall k \in \pi, s \in \pi_k, \]  

(1.6a)

**Pay period:**

\[ z_{ks} + \sum_{k \in \pi, s \in \pi_g} z_{kg} + \sum_{g \in \pi, s \in \pi_g} t_{kg}^W + \sum_{g \in \pi, k \in \pi_g} t_{gs}^j \geq 1. v_{ks}, \forall k \in \pi, s \in \pi_k, \]  

(1.6b)

**No Negotiations**

\[ x_{ks} + \sum_{i \in \pi, s \in \pi_k} x_{is}^M + \sum_{k \in \pi, j \in \pi_k} x_{kj}^M + \sum_{g \in \pi, s \in \pi_g} x_{gs} + \sum_{g \in \pi, k \in \pi_g} y_{gs} \geq (1 - d_{ks} - v_{ks}), \quad \forall k \in \pi, s \in \pi_k, \]  

(1.7a)

\[ z_{ks} + \sum_{i \in \pi, s \in \pi_k} z_{is}^M + \sum_{k \in \pi, j \in \pi_k} z_{kj}^M + \sum_{g \in \pi, s \in \pi_g} z_{gs} + \sum_{g \in \pi, k \in \pi_g} t_{kg}^W + \sum_{g \in \pi, k \in \pi_g} h_{kg}^W + \sum_{g \in \pi, k \in \pi_g} h_{gs}^j \geq (1 - d_{ks} - v_{ks}), \forall k \in \pi, s \in \pi_k, \]  

(1.7b)
Forwards Negotiations on directed Arcs

Pay period:
\[
f^W_{ks} + \sum_{g \in \succ k, k \notin g} f^W_{kg} + \sum_{g \in \succ k, k \notin g} f^J_{gs} + \sum_{g \in \succ k, k \notin g} f^W_{gs} + \sum_{g \in \succ k, k \notin g} f^J_{kg} + \sum_{g \in \succ k, k \notin g} x_{kg} + \\
\sum_{k \notin g, g \in k} x_{gs} + \sum_{k \notin g, g \in k} x^M_{ks} + \sum_{k \notin g, g \in k} x^M_{gs} + 1.\text{forward}_{ks}, \forall s \in \succ k, k \notin s,
\]

(1.8a)

\[
f^J_{ks} + \sum_{g \in \succ k, k \notin g} f^W_{kg} + \sum_{g \in \succ k, k \notin g} f^J_{gs} + \sum_{g \in \succ k, k \notin g} f^W_{gs} + \sum_{g \in \succ k, k \notin g} f^J_{kg} + \sum_{g \in \succ k, k \notin g} x_{kg} + \\
\sum_{k \notin g, g \in k} x_{gs} + \sum_{k \notin g, g \in k} x^M_{ks} + \sum_{k \notin g, g \in k} x^M_{gs} + 1.\text{forward}_{ks}, \forall s \in \succ k, k \notin s,
\]

(1.8b)

Reward period:
\[
h^W_{ks} + \sum_{g \in \succ k, k \notin g} h^W_{kg} + \sum_{g \in \succ k, k \notin g} h^J_{gs} + \sum_{g \in \succ k, k \notin g} h^W_{gs} + \sum_{g \in \succ k, k \notin g} h^J_{kg} + \sum_{g \in \succ k, k \notin g} z_{kg} + \\
\sum_{k \notin g, g \in k} z_{gs} + \sum_{k \notin g, g \in k} z^M_{ks} + \sum_{k \notin g, g \in k} z^M_{gs} + 1.\text{forward}_{ks}, \forall s \in \succ k, k \notin s,
\]

(1.9a)
\[ h_{ks}^J + \sum_{g \in \succ k, k' \succ g \atop \pi^k_g < \pi^k_{k'}} h_{kg}^W + \sum_{g \in \succ k, k' \succ g \atop \pi^k_g < \pi^k_{k'}} h_{gs}^J + \sum_{g \in \succ k, k' \succ g \atop \pi^k_g < \pi^k_{k'}} h_{gs}^W + \sum_{g \in \succ k, k' \succ g \atop \pi^k_g < \pi^k_{k'}} h_{kg}^J + \sum_{k \in \succ g \atop \pi^k_g < \pi^k_{k'}} z_{kg} + \]
\[ \sum_{k \in \succ g \atop \pi^k_g < \pi^k_{k'}} z_{gs} + \sum_{k \in \succ g \atop \pi^k_g < \pi^k_{k'}} z_{M_{kg}} + \sum_{k \in \succ g \atop \pi^k_g < \pi^k_{k'}} z_{M_{gs}} \geq 1. \text{forward}_{ks}, \forall s \in \succ k, k' \succ s, \quad (1.9b) \]

**Reward period:**

**Forward Negotiation constraints**

\[ \sum_{i \in \succ k, k \in \succ i} t_{ik}^W \leq \sum_{i \in \succ j, j \in \succ i} f_{ij}^J, \forall j \in J, \quad (1.10a) \]

\[ \sum_{i \in \succ k, k \in \succ i} z_{ik}^W \leq \sum_{i \in \succ j, j \in \succ i} f_{ij}^J, \forall j \in J, \quad (1.10b) \]

\[ \sum_{k \in \succ j, j \in \succ k} t_{kj}^I \leq \sum_{i \in \succ j, j \in \succ i} f_{ij}^W, \forall i \in P, \quad (1.10c) \]

\[ \sum_{k \in \succ j, j \in \succ k} z_{kj}^J \leq \sum_{i \in \succ j, j \in \succ i} f_{ij}^W, \forall i \in P, \quad (1.10d) \]

**Backward Negotiation constraints**

\[ \sum_{i \in \succ k, k \in \succ i} y_{ik}^W \leq \sum_{j \in \succ i, i \in \succ j} h_{ij}^J, \forall j \in J, \quad (1.10e) \]

\[ \sum_{i \in \succ k, k \in \succ i} z_{ik}^W \leq \sum_{j \in \succ i, i \in \succ j} h_{ij}^J, \forall j \in J, \quad (1.10f) \]

\[ \sum_{k \in \succ j, j \in \succ k} y_{kj}^J \leq \sum_{i \in \succ j, j \in \succ i} h_{ij}^W, \forall i \in P, \quad (1.10g) \]

\[ \sum_{k \in \succ j, j \in \succ k} z_{kj}^J \leq \sum_{i \in \succ j, j \in \succ i} h_{ij}^W, \forall i \in P, \quad (1.10h) \]
Aggregate arcs for easier use:

\[ M. z_{ij}^M \geq Z_{ij}^f + Z_{ij}^w, \forall j \in \succ_i, i \in \succ_j, \]  
\[ z_{ij}^M \leq Z_{ij}^f + Z_{ij}^w, \forall j \in \succ_i, i \in \succ_j, \]  
\[ M. x_{ij}^M \geq X_{ij}^f + X_{ij}^w, \forall j \in \succ_i, i \in \succ_j, \]  
\[ x_{ij}^M \leq X_{ij}^f + X_{ij}^w, \forall j \in \succ_i, i \in \succ_j, \]  
\[ M. y_{ij}^M \geq Y_{ij}^f + Y_{ij}^w, \forall j \in \succ_i, i \in \succ_j, \]  
\[ y_{ij}^M \leq Y_{ij}^f + Y_{ij}^w, \forall j \in \succ_i, i \in \succ_j, \]  
\[ M. t_{ij}^M \geq T_{ij}^f + T_{ij}^w, \forall j \in \succ_i, i \in \succ_j, \]  
\[ t_{ij}^M \leq T_{ij}^f + T_{ij}^w, \forall j \in \succ_i, i \in \succ_j, \]

Keep track of forward and backward negotiations for un-directed arcs

\[ M. d_{ij}^a \geq \sum_{k \in \succ_i, j \notin \succ_k} t_{ik}^w + \sum_{i \in \succ_j, j \notin \succ_i} t_{ij}^f + \sum_{s \in \succ_j, j \notin \succ_s} z_{sj}^w + \sum_{i \in \succ_j, j \notin \succ_i} z_{ji}^f \forall i \in \succ_j, j \in \succ_i, \]  
\[ M. v_{ij}^a \geq \sum_{k \in \succ_i, j \notin \succ_k} y_{ik}^w + \sum_{i \in \succ_j, j \notin \succ_i} y_{ij}^f + \sum_{s \in \succ_j, j \notin \succ_s} x_{sj}^w + \sum_{i \in \succ_j, j \notin \succ_i} x_{ji}^f \forall i \in \succ_j, j \in \succ_i, \]
\[ d_{ij}^u \leq \sum_{k \in i, i \not\in k} t_{ik}^W + \sum_{l \in j, j \not\in l} t_{lj}^J + \sum_{s \in j, j \not\in s} z_{sj}^W + \sum_{i \in g, g \not\in j} z_{ig}^J \quad \forall i \in j, j \not\in i, \quad (1.12c) \]

\[ d_{ij}^w \leq \sum_{k \in i, i \not\in k} y_{ik}^W + \sum_{l \in j, j \not\in l} y_{lj}^J + \sum_{s \in j, j \not\in s} x_{sj}^W + \sum_{i \in g, g \not\in j} x_{ig}^J \quad \forall i \in j, j \not\in i, \quad (1.12d) \]

Keep track of forward and backward negotiations for directed arcs

\[ M.d_{ij}^d \geq \sum_{k \in i, i \not\in k} t_{ik}^W + \sum_{l \in j, j \not\in l} t_{lj}^J + \sum_{s \in j, j \not\in s} z_{sj}^W + \sum_{i \in g, g \not\in j} z_{ig}^J, \]

\[ \forall j \in i, i \not\in j \land \forall i \in j, j \not\in i, \quad (1.12e) \]

\[ M.v_{ij}^u \geq \sum_{k \in i, i \not\in k} y_{ik}^W + \sum_{l \in j, j \not\in l} y_{lj}^J + \sum_{s \in j, j \not\in s} x_{sj}^W + \sum_{i \in g, g \not\in j} x_{ig}^J, \]

\[ \forall j \in i, i \not\in j \land \forall i \in j, j \not\in i, \quad (1.12f) \]

\[ d_{ij}^u \leq \sum_{k \in i, i \not\in k} t_{ik}^W + \sum_{l \in j, j \not\in l} t_{lj}^J + \sum_{s \in j, j \not\in s} z_{sj}^W + \sum_{i \in g, g \not\in j} z_{ig}^J, \]

\[ \forall j \in i, i \not\in j \land \forall i \in j, j \not\in i, \quad (1.12g) \]

\[ v_{ij}^w \leq \sum_{k \in i, i \not\in k} y_{ik}^W + \sum_{l \in j, j \not\in l} y_{lj}^J + \sum_{s \in j, j \not\in s} x_{sj}^W + \sum_{i \in g, g \not\in j} x_{ig}^J, \]

\[ \forall j \in i, i \not\in j \land \forall i \in j, j \not\in i, \quad (1.12h) \]

1.5.6 Cycle-Based Approximation Solution for Perfect Relaxed-Stable Matching

1.5.6.1 Pre-processing

DA algorithm and deserving agents. As explained in section (5.1) people-proposing and job-proposing DA algorithms can result in different match sizes as preferences can be one-sided acceptable meaning that agents’ preferences are not
consistent. Let us set our milestone DA stable matching to the solution with greater match cardinality. From unmatched agents of DA matching, we find the set of deserving agents. Remember that deserving agents are those who are unmatched under DA matching and have greater frequency than the assignment of any of their acceptable items.

Centrality measures. Before any algorithmic attempt, we introduce a centrality measure for each agent in the market indicating the compatibility of an agents’ preferences to market. Greater value of this measure is associated with higher likelihood for agent being matched under any stable matching.

The nature of constructing this measure categorizes it under the wide umbrella of eigenvector centrality measures. Let $MMEC_i$ denote the Matching Market eigenvector Centrality measure for any agent $i \in P \cup J$.

If $b_{ij}$ is equal to 1 when agent $i$ has $j$ in its acceptable list and $\pi_j^i$ denotes the ranking of agent $i$ in $j$’s list, and $K$ is a constant number denoting the maximum number of preferences listed, then MMEC measure is calculated through the following formula:

$$MMEC_i = \sum_{j \in \succ_i} \left(1 - \left(\frac{\pi_j^i - 1}{n}\right)\right)\left(\frac{b_{ij}}{K - |\succ_j|}\right) \quad \forall \ i \in P \cup J$$

$MMEC_i$ is constructed from two components. The first component examines how agent $i$ is ranked relative to all agents in the opposite side. The other component focuses on the rank of $i$ in comparison to those acceptable to agent $j$. $MMEC_i$ lies between zero and $n$ ($MMEC_i \in [0, n]$). The minimum value happens when agent $i$ is not accepted by any agent in the opposite side, the maximum value happens when agent $i$ is accepted by everyone in the opposite side and is ranked first in an acceptable list of length one.
It is important since the number nonzero parts of this summation is counted through the binary $b^i_j$ variable, then a combined centrality measure, denoted by $CC_i$, indicates the likelihood of an agent getting matched under any stable matching. It can also be interpreted as how much agent $i$ is listed in the best ranked positions of its preferred options.

Let this combined measure be formulated as below:

\[
CC_i = \begin{cases} 
\frac{MMEC_i}{\sum_{j \in |\succ_i|} b^i_j}, & \text{if } \sum_{j \in |\succ_i|} b^i_j > 0 \\
0, & \text{otherwise}
\end{cases}
\]

Although these measures do not provide a rigid threshold that separate agents that remain unmatched under a stable matching, but they still provide valuable insights on how market treats agents.

**Proposition.** Agents with $CC_i$ equal to zero are unmatched agents under any stable matching.

**Observation.** Agents with lower $CC_i$ are associated to unmatched agents under any stable matching.

By construction, $CC_i$ is designed to show the likelihood of an agent getting matched under any stable matching. therefore lower values show smaller probability for agents to be assigned in a stable matching.

**Insight.** Prioritizing matches based on lower $CC$ values can increase match cardinality.

If unmatched agents under DA algorithm who are associated to low $CC_i$ are matched, then there are still a great probability for agents who are matched under DA to have a matching. as a result, prioritizing low $CC_i$ values in the process of matching can increase match cardinality.
1.5.6.2 Algorithmic method:

This section involves iteratively generating a network from agents and their acceptable items. Each arc of the network represents a potential match. The assignment of every agent is finalized once it is matched in both periods. The main idea is to find cycles with each arc alternately spanning over different periods. Every node in a cycle has an incoming and an outgoing arc alternating over different periods by construction. The advantage of a cycle is that, once it is accepted, it finds an assignment of every agent in the cycle for both periods.

Generating graph:

Iteration 1: every person $i \in P$ applies to its most favorite job $j$ in its list. This application is represented by an arc in the network from $i$ to $j$. Then job $j$ applies to people that it prefers more than person $i$ in its acceptable list. If there is a person $g$ with greater frequency than person $i$ who has job $j$ in its acceptable list, then person $g$ applies to job $j$ as well, represented by an arc from $g$ to $j$ in the network.

At any iteration, the cycle-based heuristic finds a set of matches that are finalized. We remove any node that is matched from graph of iteration $k - 1$ and all their associated arcs. Then we start the next iteration (iteration $k$) of generating graph.

Iteration $k \geq 2$: every agent $i$ applies to its most favorite item $j$ that is not yet proposed to by $i$. This application is represented by an arc in the network. Then agent $j$ applies to all items that are better ranked than $i$ in its acceptable list. In addition, if there is an agent $g$, in the same set as agent $i$, who has greater frequency than $i$ and lists $j$ in its acceptable list, then agent $g$ applies to agent $j$ as well, represented by an arc from $g$ to $j$ in the network.
aforementioned three steps are applied for all nodes in both sets. The resulting network is copied for the second period network. Note that the resulting graph is a directed graph $G(4n, A)$, with $4n$ nodes, where $n$ is the number of agents in each set, and $|A|$ arcs.

**Finding cycles in graph.** For every agent $i \in P$, find the set of candid cycles $C$ with each arc alternating over opposite periods and potentially assigning each person (job) to a job (person) that is at least as preferred as its own-proposing DA matching in one period. If cycle $c \in C$ includes a deserving agent, then it fulfills the negotiation condition and is called negotiation cycle. In such cycles stability of one period is relaxed for a deserving node, only if there is a reward for its match in the next period. Since the graph is symmetric for both periods and by construction graph is bipartite, there is no need to find cycles starting from agent in set $J$ too.

**Accepting a cycle.** The third step involves choosing between the set of candid cycles. The cycle choosing criteria consists of two parts: 1) Utility measure that calculates the utility gained by each agent when matched in two periods, and 2) negotiation measure that adds the benefit captured in a negotiation. This criteria is the summation of utility and negotiation measures divided by cycle length.

If $u_i^j$ be the utility gained by agent $i$ if matched to node $j$ and $DN_c$ be the deserving nodes in cycle $c$, then $\Pi_c$ shows the utility gained by all potential matches in cycle $c$ and $N_c$ shows the benefit gained from all potential negotiations in cycle $c$.

$$
\Pi_c = \sum_{i \in c} \left( u_i^{i+1} + u_{i+1}^i \right), \quad N_c = \sum_{i \in DN_c} \left( u_i^{i+1} + F_i \right)
$$

$$
Criteria_c = \frac{\Pi_c + N_c}{|c|}
$$

43
After finding the criteria value associated with each cycle, we sort cycles in descending order of their criteria value. We sort cycles with equal criteria based on their total $CC$ value in ascending order. Total $CC$ value of a cycle is simply the summation of $CC_i$ for all nodes in cycle $c$. Once a cycle is accepted, other cycles with same agents are eliminated from the set of candid cycles $C$.

Sorting cycles by criteria prioritizes those with greater average utility that is aligned with the objective function of our proposed integer programming model. Prioritizing low $CC$ cycles between those with equal criteria helps matching agents that are likely to remain unmatched. Due to the construction of our cycles, accepted matches are not going to be blocked by their DA match in both periods.

**Aftermarket.** Since acceptable lists are incomplete, there will be a point where no cycles can be found in the graph after all agents are added to the network. We find paths that go through both periods alternately and potentially assign each person to a job that is at least as preferred as his/her DA matching. Similarly, we find paths that go through both periods alternately and potentially assign each job to a person at least as preferred as its DA matching.

After finding the criteria value associated to each path, we sort paths in descending order based on their criteria. We prioritize paths that start with the tail of the last accepted paths. Paths with equal criteria are sorted based on their total $CC$ value in ascending order. Once there are no paths going through both periods, we accept paths of length one that are over only one period with the same accepting process. That is arcs that represent one-sided acceptable matches are accepted at the end of the procedure. If there exit an agent that is not matched at the end of the heuristic, that means it was not chosen in utility maximization or negotiations.
paths. Such nodes are agents that remain unmatched under DA matching. In addition, accepted negotiations uncured greater value than the negotiations offered to match the remaining unmatched nodes.

1.6 Computational Study

Let $Pr_p$ denote the probability that person $p$ has job $j$ in its acceptable preference list, and $Pr_j$ denote the probability that job $j$ has person $p$ in its acceptable preference list. We implemented our proposed mechanism on random Erdos-reyni graphs $G(4n, Pr_p, Pr_j)$ and also on correlated graphs where each personnel order their acceptable jobs with 90% correlation and vice versa. The results from our proposed Integer programming model and heuristic is compared with the deferred acceptance algorithm. Comparison with DA algorithm serves as a milestone that points out the contribution and social impact of our mechanisms. Table 1. includes measures used to compare the performance of the three methods. Table 1 compares the performance of the three methods when acceptable lists are randomly ordered and Table 2. compares the performance of the three methods when acceptable lists are ordered with a 90% correlation ratio.

Table 1.1: Network scenarios for computational study

<table>
<thead>
<tr>
<th>Network setting</th>
<th>Network size $(4n)$</th>
<th>Network structure and $Pr_p$, $Pr_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Graphs</td>
<td>80</td>
<td>Equal Probabilities 0.4, 0.4</td>
</tr>
<tr>
<td>Correlated Graphs</td>
<td>200</td>
<td>Equal Probabilities 0.7, 0.7</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>Unequal Probabilities 0.2, 0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unequal Probabilities 0.3, 0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unequal Probabilities 0.5, 0.8</td>
</tr>
</tbody>
</table>

45
The results from preliminary runs show that proposed heuristic reduced the number of people who remain unsatisfied by 84% and 81% with random and correlated preferences respectively using the proposed heuristic. Under the proposed heuristic, on average only 0.02% of total number of people and jobs end up unsatisfied in both matching cycles across different network structure scenarios. This percentage is 0.03% for markets with correlated preferences. On the other hand, when using DA algorithm, the expected number of people and jobs that end up unsatisfied in both matching cycles is 14% for markets with random preferences and 18% for markets with correlated preferences.

Number of negotiations indicates the number of times stability is relaxed. In order to explain the heuristic mechanism we compare the number of deserving nodes with the number of deserving nodes that are matched with the help of the negotiations. When preferences are random 77% of the deserving nodes are matched through the proposed mechanism and while correlated preferences results in 94% of deserving nodes getting matched.

Figure 1 and 2 show the trajectory of unsatisfied agents across six network scenarios shown in Table 1 when the network size equals to 80. when preferences are long DA algorithm is able to find solutions that have few unmatched agents, however as preferences are expected to have shorter length the number, there is a significant difference between DA algorithm and proposed mechanisms.

1.7 Conclusion

Our results indicate that large-scale two-sided markets significantly benefit from the matching with negotiation mechanism. In particular, markets where information
is limited, agents are not aware of all of their options or making applications are costly, incomplete preference lists are popular. Our proposed mechanism, on average, matched 73% of unmatched agents under DA algorithm in markets with random preferences and 78% of unmatched agents under DA algorithm in markets with correlated preferences. Satisfying people and jobs that would remain unmatched normally has drastically improves social welfare of all agents in the market over two periods.
Figure 1-3: Trajectory of number of unsatisfied agents in both periods for markets with random ordering of preferences in acceptable items using Deferred Acceptance algorithm (DA), Integer Programming (IP) model and Heuristic (H)
<table>
<thead>
<tr>
<th></th>
<th>$G(80, Pr_p, Pr_j, CR = 0%)$</th>
<th>(.2,4)</th>
<th>(.4,4)</th>
<th>(.4,.5)</th>
<th>(.4,.7)</th>
<th>(.5,.8)</th>
<th>(.7,.7)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DA</strong></td>
<td>Unmatched in job-proposing</td>
<td>32</td>
<td>13</td>
<td>12</td>
<td>5</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Unmatched in people-proposing</td>
<td>32</td>
<td>12</td>
<td>12</td>
<td>5</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Minimum unsatisfied in both</td>
<td>32</td>
<td>12</td>
<td>12</td>
<td>5</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>periods</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Negotiations</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>IP</strong></td>
<td>Unmatched</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Unsatisfied in both periods</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Negotiations</td>
<td>5</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>H</strong></td>
<td>Unmatched</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Unsatisfied in both periods</td>
<td>6</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Negotiation cycles</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Deserving Nodes</td>
<td>20</td>
<td>12</td>
<td>11</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Deserving Nodes Matched</td>
<td>17</td>
<td>5</td>
<td>9</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 1.3: Comparison of Deferred Acceptance algorithm (DA), Integer Programming (IP) model and Heuristic (H) for markets with 90% correlation in preference ordering of acceptable items

<table>
<thead>
<tr>
<th></th>
<th>G(80, Pr_p, Pr_j, CR = 0.90%)</th>
<th>(.2,.4)</th>
<th>(.4,.4)</th>
<th>(.3,.5)</th>
<th>(.4,.7)</th>
<th>(.5,.8)</th>
<th>(.7,.7)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unmatched in job-proposing</td>
<td>36</td>
<td>16</td>
<td>20</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Unmatched in people-proposing</td>
<td>40</td>
<td>16</td>
<td>20</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Minimum unsatisfied in both</td>
<td>36</td>
<td>16</td>
<td>20</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>periods</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Negotiations</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DA</td>
<td>Unmatched</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Unsatisfied in both periods</td>
<td>6</td>
<td>2</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Negotiations</td>
<td>9</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>IP</td>
<td>Unmatched</td>
<td>6</td>
<td>4</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Unsatisfied in both periods</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Negotiation cycles</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Deserving Nodes</td>
<td>20</td>
<td>10</td>
<td>12</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Deserving Nodes Matched</td>
<td>16</td>
<td>10</td>
<td>10</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
Chapter 2

Integrating Supply Chains for Emergencies and Ongoing Operations in UNHCR

2.1 Introduction and Motivation

Improving network design to cut cost and reduce response time is critical for humanitarian logistics [87, 88]. The trend among larger humanitarian organizations (HOs) such as International Federation of Red Cross and Red Crescent Societies (IFRC), United Nations High Commissioner for Refugees (UNHCR), World Food Program (WFP), Cooperative for Assistance and Relief Everywhere (CARE), and United Nations Children’s Fund (UNICEF) is to preposition un-consigned relief items for emergency response in warehouses located close to disaster-prone areas (see e.g., [46, 78, 49, 59, 57, 9] McCoy and Brandeu, 2011; Bemley et al., 2013; Komrska et al., 2013). However, stock prepositioning is expensive, and owing to funding restrictions,
other alternatives have been suggested including vendor-managed inventory (S8), framework agreements (6), and transfer mechanisms between programs (13). Most HOs are engaged in both long- and short-term (emergency relief) operations, for which they usually operate different supply chains with separate warehouses. Recently, joint supply chains with vehicles serving both types of operations have been suggested as an alternative for saving cost (11 83). In principle, one should be able to integrate the two to reduce response time and total operating cost. Long-term operations could, for instance, be serviced by using un-consigned stock to avoid long lead times from distant suppliers, while emergency relief could be serviced by using closely located stockpiles to avoid expensive express shipments over long distances. Integration of the two supply chains may allow for additional warehouses closer to demand points, but doing so would require designing and operating a joint network with different demand uncertainties, objectives, and operational procedures. In the current paper, we attempt to address those challenges. To the best of our knowledge, models that explicitly combine emergency relief and longer-term operations supply chains for prepositioning of goods have not been reported. Furthermore, factors related to political and security conditions, often mentioned in extant literature as being important for location decisions in the humanitarian context, lack empirical or quantitative justifications. The objective of this study is to develop a model that integrates factors such as hardship, security, pilferage, co-location, and accessibility in determining best joint prepositioning warehouse locations. Such a model can help quantify the impact of an expanded network on both lead time and cost. Using the developed model, we generate empirical and computational insights on the factors that significantly influence warehouse location choice. To address the above objectives, we conducted an in-depth, exploratory case study with UNHCR, which is mandated to lead and coordinate international operations to safeguard the
rights and well-being of refugees and resolve refugee problems worldwide. This case demonstrates relevant examples of the generic challenges faced by HOs in terms of reducing costs while maintaining response in a timely manner in both short- and long-term operations. As is typical for a HO, UNHCR runs two supply chains. First, in their emergency relief (ER) supply chain, speed is essential, and they have responded to this need by using fast means of transportation from large centralized global warehouses. Second, and in parallel, they support their long-term ongoing operations (OOs), for example, camp operation, with stocks transported from de-centralized country warehouses or shipped directly from suppliers. Merging of the two supply chains presents an opportunity to reduce cost but also implies the need to redesign their network, particularly considering the locations of their global warehouses. One of the challenges associated with building a joint supply chain constitute the differing objectives, with ER aiming to reach the beneficiaries as fast as possible, and OOs aiming to satisfy all demand incurring minimum cost, while still being mindful of responsiveness. Hence, decisions have to consider both objectives and their trade-offs, requiring multi-objective models. In this study, we developed such a model, tested it with datasets based on the UNHCR case study, and presented an efficient frontier analysis for different budget levels. UNHCR is now working on implementing the model into their enterprise resource planning (ERP) system and as a part of their overall decision making process. This paper answers the call for more applied, context-sensitive research in humanitarian operations\textsuperscript{[43]} with two main research contributions. First, we developed a framework and a data-driven model that integrate long-term and emergency relief supply chains, accounting for both response time and cost. Second, the study helps fill the gap in humanitarian network design literature by including security and political factors, which influence warehouse locations for prepositioned stock but have not been incorporated in decision
models thus far. Through computations, we show their impact on transportation and warehousing decisions.

2.2 Literature review

Humanitarian research typically categorizes network design models depending on problem type, objective functions, number of levels included, and the manner in which uncertainties are treated ([18, 17, 44, 68, 74, 69]). In their overview of models, Caunhye et al. (2012) divided the literature into two main categories, namely, facility location and relief distribution. Our research falls under the former category, which can be subdivided into pure location models (e.g., [18]), inventory models (e.g., [16]) and models, which, similar to our approach, combine the location problem with the amount of inventory to preposition at each location (e.g., [57]). The objective for models addressing public sector problems and humanitarian aid is often to maximize the amount of demand satisfied or to minimize lead time (e.g., [7, 61, 77, 15, 27, 74]). Comprehensive reviews were presented in Caunhye et al. (2012), Holguin-Veras et al. (2013), Rennemo et al. (2014), and Uzdamar and Ertem (2015) as well. To ensure maximum demand coverage, some researchers have associated a penalty cost with unmet demand and endeavored to minimize said cost (e.g., [73, 72]). Another stream of research has focused on modeling responsiveness and has generated models that include a benefit for timely demand satisfaction in their objective function along with a penalty for unmet demand (review in [44]). In contrast, other studies have focused primarily on the cost aspect when solving the facility location problem (FLP) in the humanitarian aid context. For example, Iakovou et al. (1997) considered only facility and operating costs, while Barcinpour and Esmaeili (2014) and Bozorgi-Amiri et al. (2013) accounted for several cost considerations as a joint
objective function. In this paper, we include both lead time and cost as objectives, along with other considerations such as security and political factors, while enforcing total demand satisfaction. For considering multiple objectives, one common technique used is to associate weights with each objective \( [25,67] \) and solve the FLP as a cost minimization problem (e.g., \([79]\)). Because providing a single optimal solution when multiple objectives are involved remains difficult, different objective functions are often assigned weights to prioritize them, and solutions for different prioritization schemes are analyzed accordingly \( [90]\). One example is Tzeng et al. (2007), wherein fuzzy logic was applied to find the optimal balance among cost, travel time, and proportion of demand met. Stauffer et al. (2015), in contrast, modeled responsiveness by including an expediting cost in the objective function. Another alternative to handle multi-objectives when the main consideration is to maximize responsiveness is including additional objectives such as cost as model constraints (e.g., \([77,27]\)). We adopted this approach and modeled responsiveness by minimizing lead time as a measure of service and included a budget constraint. Our model was then used to determine \((1+\epsilon)-\text{Pareto curves}\) to analyze the trade-off between the cost and lead time objectives. Such methods are commonly used in multi-objective optimization \( [70]\). Two studies are of particular interest to our research. Stauffer et al. (2015) looked at how organizations can manage short- and long-term programs together. They found that a centralized hub configuration combined with temporary hubs can reduce overall supply chain costs over a long time horizon when global vehicle supply chains for development and emergency operations are handled together. Similarly, Besiou et al. (2014) concluded that using the same vehicles for both development and emergencies affects performance ranking for service levels, suggesting that organizations “should include the operational mix in the decision-making factors when choosing the structure” (ibid, p.10). Accordingly, both aforementioned
studies demonstrate the advantages of integrated network structures. The literature on commercial FLPs ([60]) lists several important factors from the network design viewpoint. We identified additional factors by reviewing humanitarian logistics literature and modeling papers (Appendix A). We focused on the extent to which extant research discusses, models, and quantifies factors based on real data. We found that similar to commercial models, most humanitarian models include infrastructure, demand, and cost, but there are some differences. In humanitarian research, greater weight is given to demand risk ([55] [74]). Funding issues are typically modeled as budgetary constraints ([77]). Furthermore, political and security issues are accorded greater concern ([56]). While various papers have discussed these factors, modeling has been limited to demand and logistics, except for one security aspect, namely, personnel availability ([77] [15]). Duran et al. (2011), for example, have provided the most extensive discussion of political and security factors, but rather than quantifying and including them in the model, they suggested ranking the locations resulting from the quantitative analysis, which requires decision makers to qualitatively judge each location based on how well it meets these additional factors. Accordingly, there is a big gap in prevailing research regarding (i) empirical analysis of the extent to which each factor actually impacts location decisions and (ii) quantification of any factor’s potential impact using real data and its incorporation into the model. Our study contributes to filling this gap.
2.3 Methodology for qualitative empirical study

2.3.1 Case Selection

An in-depth, exploratory case study was conducted with UNHCR following theory-building principles ([29, 58, 63, 92]). A single case allows one to gain more in-depth understanding of the studied phenomenon ([91]). Accordingly, the UNHCR case provides an opportunity to acquire the rich, qualitative data required to develop a network design model useful for the humanitarian sector. The study was initiated in response to a 2012 request from UNHCR for help with developing a tool to support their warehouse location decisions. UNHCR’s effort to reduce both cost and response time in long- and short-term operations is representative of the efforts of most large HOs. Many organizations including WFP and UNICEF operate in the same countries as UNHCR under comparable political and security conditions, and have similar organizational structures and human resources policies. The case study included a field trip to UNHCR’s Kenya operations ([10]). This operation represents a critical case ([71]) considering that: i) it is among the top 20 UNHCR operations; ii) it has a global warehouse and multiple operations in different parts of the country; iii) it has a combination of local and international procurement and deliveries, and iv) in the context of external validity ([29, 58]), UNHCR regards its Kenyan operations to be representative of other large-scale operations such as in Syria, Pakistan, and Afghanistan.

2.3.2 Case Description

With more than 8,600 staff and an annual budget of about USD 4.3 billion, UNHCR helps approximately 34 million people each year in more than 125 countries
The largest operations constitute assistance to refugees and internally displaced people in Asia (e.g., Afghanistan), the Middle East (e.g., the Syrian crisis), and Africa (e.g., Kenya, Democratic Republic of the Congo, and South Sudan). As a sector leader in emergency shelter and camp management, a large portion of UNHCR’s work is in the form of emergency response. Moreover, UNHCR operates camps on a longer-term basis. One example is the 20-year-old refugee camp in Dadaab, Kenya (www.unhcr.org).

To support these operations, the organization depends on a three-level network structure, represented by: i) global warehouses; ii) country warehouses; and iii) supplier ship-out locations. UNHCR operates seven global warehouses, which stock core relief items (CRI) and specialized items such as information technology (IT) equipment and vehicles. The CRIs include tents, tarpaulins, mosquito nets, blankets, sleeping mats, plastic buckets, jerry cans, kitchen sets, and solar lanterns. Global stock is un-consigned, meaning that the items are not dedicated to a certain country’s operations, but stored in bonded facilities or imported under duty-exempt status. In contrast, country warehouses hold consigned, customs-cleared stock as a buffer for multiple distribution points within national borders. UNHCR uses its country warehouses as merging points, combining internationally delivered items and locally purchased goods before sending them to the relevant distribution points. The organization reduced the number of country warehouses from 350 in 2013 to 192 in 2014. UNHCR procures its CRIs from a large number of suppliers. Based on 2013 statistics, the top countries for procurement of stored items include Belgium, China, Denmark, India, Jordan, Kenya, Kuwait, Lebanon, Pakistan, Turkey, Syria, and the United Arab Emirates. At present, only a small fraction of the stored items is sent directly from the suppliers to the distribution points, and one of this study’s con-
tributions is to help UNHCR determine whether more direct methods of shipment from the suppliers to the distribution points would be beneficial.
<table>
<thead>
<tr>
<th>Step</th>
<th>Date</th>
<th>Method (#/length)</th>
<th>Purpose</th>
<th>Interviewee/source</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>May-Sept, 2012</td>
<td>Three exploratory interviews (1 h each)</td>
<td>Planning and designing study, including field trip</td>
<td>Senior Business Analyst, Head of Logistics &amp; Operations; Head of Supply Management Service Officer</td>
<td>HQ in Budapest</td>
</tr>
<tr>
<td>2</td>
<td>Oct 14–16, 2012</td>
<td>Two semi-structured interviews (1 h each); One group discussion (half day)</td>
<td>Understanding how UNHCR has set up its network of warehouses and manages operations; Collecting qualitative data for designing the model</td>
<td>Supply Officer, Assistant Supply Officer</td>
<td>Two of the four distribution points in Dadaab, Kenya</td>
</tr>
<tr>
<td></td>
<td>Oct 17–19, 2012</td>
<td>Three semi-structured interviews (1 h each); One group discussion (half day)</td>
<td></td>
<td>Senior Supply Officer, Assistant Supply Officer, Assistant Program Representative, Warehouse manager (Kuehne + Nagel)</td>
<td>One global warehouse and one country warehouse located in Nairobi, Kenya</td>
</tr>
<tr>
<td>3</td>
<td>Nov 18, 2012</td>
<td>Questionnaire (building on the list of factors in Appendix A)</td>
<td>Identifying contextual factors important to consider for the model</td>
<td>Senior Business Analyst, Head of Logistics &amp; Operations, Senior Supply Officer</td>
<td>HQ in Budapest</td>
</tr>
<tr>
<td>4</td>
<td>Oct-Nov, 2012</td>
<td>ERP system data</td>
<td>Gathering data, including information about core relief items, warehouse points, demand points, goods flow, and transportation costs</td>
<td>UNHCR's global ERP system, Information Management Officer</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Sept-Oct, 2014</td>
<td>Multiple structured interviews (on a weekly/daily basis)</td>
<td>Designing and validating the network design model for use in a humanitarian context</td>
<td>Senior Business Analyst, Head of Logistics &amp; Operations</td>
<td>HQ in Budapest (via Skype and mail)</td>
</tr>
<tr>
<td>6</td>
<td>Nov-Dec, 2014</td>
<td>ERP system data, internet search on security levels, personnel, pilferage, disaster scenarios</td>
<td>Gathering and populating OO and ER demand, cost, and lead time data followed by cleaning, estimation, and discussion—for details, see Section 4.4 and Appendix B.</td>
<td>UNHCR Management Service, Head of Inventory &amp; Warehouse Management Unit, Chief of Emergency Coordination Unit</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Jan-June, 2015</td>
<td>ERP system data, internet search for relevant data</td>
<td>Quantifying identified contextual factors and incorporating them in the developed model (cf. Section 4.4)</td>
<td>UNHCR's global ERP system, UNHCR supply chain management team, Brambles, ongoing consultancy project at UNHCR</td>
<td>HQ in Budapest (via Skype and mail)</td>
</tr>
</tbody>
</table>
2.3.3 Data Collection

By combining a range of qualitative and quantitative methods, and by using several data sources and investigators, we aimed to triangulate the collected information ([71]) and increase internal validity ([29, 58, 63]). Data was collected in seven steps (Table 2.1). Throughout data collection, the case study protocol was updated, including research instruments and procedures, interviewee details, interview transcriptions, field notes, case summaries, and preliminary findings (Yin, 2014). All collected data were summarized and sent back to the respondents, who then commented and confirmed the findings.

2.3.4 Data Analysis

For data analysis, we followed the analytical abstraction process of Miles and Huberman (1994) by summarizing and categorizing data collected through interviews, group discussions, and questionnaires. First, the data were coded using a provi-
sional list based on the factors identified in the literature review (Appendix A). Examples included budget constraints, demand risks, and infrastructure limitations. Moreover, by using open coding ([30]), additional coding categories such as hardship, security, pilferage, and co-location emerged during the analysis. The resulting coding list facilitated a logical link between the collected data and the constructed model, where each of the identified factors was, as explained in Section 2.4.3, included in model development. In the second step, referred to herein as pattern or axial coding ([30] [63]), memos from all interviews and group discussions were compared with recurring phrases and threads in the questionnaire responses to identify emerging trends and themes. One such theme was the respondents’ emphasis on the difficulty of merging the two supply chains for ongoing operations and emergency response, and the different objectives of minimizing cost and minimizing response times. Additional insights from this step are discussed in Section 2.4.1 and form an important input in model development. In the final step, all findings were discussed and confirmed with key UNHCR staff to validate the inputs used for model design and analysis.

2.4 Model development based on empirical findings

This section presents the empirical findings of importance for constructing the model, starting with the background to UNHCR’s supply-chain strategy and why they want to merge the two supply chains. Then, we present the influencing factors identified in the qualitative study (4.2), the model and how it incorporates the factors (4.3), and the data used for analysis (4.4).
2.4.1 Merging Two Supply Chains

Based on an analysis of past emergency response patterns, UNHCR senior executives have, with support from major UNHCR donors, decided to install an immediate response capacity of CRIs for 600,000 beneficiaries. To enable emergency response within 72 hours, UNHCR has set up a network of global warehouses for prepositioning with fast means of transportation. However, similar to many HOs, they must consider not only time but also cost. As a critical step to lowering their total operational costs, UNHCR will merge their two supply chains. The ER-supply chain deals with highly uncertain demand occurring in sudden-onset, man-made disasters, and it is designed to minimize response time where central funds are used to buy and preposition unconsigned stock in one of the global warehouses. UNHCR reaches out centrally to a few of its big donors such as The Department for International Development in the UK (DFID) to pre-fund the central emergency stock, which can then be used to support global operations. In the case of an emergency, global stock can be “bought” by country operations and delivered by fast means of transportation as consigned stock for local consumption. The OO supply chain deals with long-term operations characterized by continuous demand and relatively low uncertainty. It is designed to minimize cost and involves decentralized consigned stock bought under a country’s budget and pushed to that country’s warehouses or items shipped directly from suppliers. Once stock is shipped and clears customs in a country, it is difficult and expensive to re-export the goods to another country. The decentralized network structure of OO originates from the historical set-up of the organization, with a weaker supply chain center at its headquarters (HQ) and strong UNHCR country operations. Country operations are managed by a UNHCR Representative, appointed by the High Commissioner, based on an agreement with the host country.
As such, the host country can influence these operations, but decisions pertaining to operational priorities and fund utilization reside with UNHCR. Country operations are in charge of their own budgets, in addition to reaching out to local donors. Similar to other large HOs such as WFP, local program managers decide how and when to spend their money depending on when it is made available from donors to country operations by a HQ-approved budget and spending authority. In other words, sourcing is driven by funding with limited supply chain focus, which has resulted in sub-optimization with excessively large stock in a few locations and very low stock in others. Considering that UNHCR, akin to many organizations in the public sector, operates on annual budgets, inventory has generally been regarded as a safety measure not “to lose money at year end,” following the mentality that “it is always good to have stocks, we might have use for it and we do not lose the money,” resulting in a high amount of dead stock.

### 2.4.2 Influencing factors identified in qualitative study

The empirical findings unearthed nine factors that can be categorized in three groups: i) demand characteristics; ii) logistics; and iii) political and security situational factors (Figure 2-1).

#### 2.4.2.1 Demand Characteristics

The first group of factors represents demand characteristics, and it is related to demand risk and budget constraints (Model A). OO demand is in general continuous with low uncertainty, and it can be estimated from historic data. ER demand is encountered in sudden-onset disasters and is highly uncertain. ER demand is difficult to predict and requires scenario planning by considering the probabilities of different
scenarios.

### 2.4.2.2 Logistics

Logistics-related factors (Appendix A) identified in this study include capacity and cost of transportation and warehousing, lead times among suppliers, warehouses, and demand points, as well as accessibility. Transportation and warehousing capacities are typically not an issue, although the price of contracting additional transportation or warehouse space may vary. In transshipment points such as ports, humanitarian cargo may have to “compete” with fluctuating commercial transport flows. Furthermore, considering that UNHCR, similar to many other HOs, does not have its own fleet for transporting relief supplies, prices typically increase in the event of a disaster ([73, 15]), and they may vary depending on the location and the disaster situation. Moving global warehouses closer to demand points could increase the cost of restocking said warehouses, depending on the means of transportation required and available for accessibility due to poor infrastructure. By placing the global warehouses closer to the demand points, the shipping cost, as well as lead time for this leg should be reduced. Physical accessibility aspects include available primary and secondary roads; sea freight and railway networks; and proximity to points of entry, including ports and airports. Without proper accessibility, the total operations cost of a warehouse site may increase, which brings with it the risk of jeopardizing the entire operation. Telecommunications infrastructure must also be in place to enable efficient communication and coordination within the organization and externally with the supply chain and implementing partners.

#### 4.2.3 Political and security situation

Building on the literature review (Appendix A), the following factors related
to political and security situation, were identified in the current study: relationship with government, security, pilferage, access to human resources (hardship), and co-location. First, HOs depend on good relationships with host governments. By signing bilateral agreements, governments can facilitate exemptions and customs-clearance procedures when goods enter or leave the country. In some cases, goods can be cleared within a day, and in other cases, the process may take weeks or months. Moreover, governments may agree to offer land or facilities at low or no cost to the organization, thus greatly reducing certain locations’ fixed costs. Second, due to its mandate, many of UNHCR’s operations are set up in the midst of political instability, including military activity and civil war. Under such circumstances, running efficient, secure logistics operations can be very difficult and may require significant security arrangements. Security concerns also include pilfering and looting. Apart from the value of the goods stolen, the very risk of theft implies increased security costs in terms of insurance and guards. Lost goods could also result in ill will for the organization due to negative media publicity, damage to donors’ image, and impeding future funding. Third, similar to most HOs, UNHCR employs both international and local staff. The ability to attract qualified workforce is a critical factor in deciding where to locate global warehouses. An important aspect is the rotation system for international personnel, which is based on hardship principles, implying that international staff spends only a limited time at a certain location. Hardship is often the toughest where need is the greatest, and rotation is extremely high. Following UN conditions of service, international employees at duty stations in high-risk areas also go on a five-day leave every four weeks. These contractual conditions mean that a greater number of personnel is required to run the operation. Fourth, in spite of competition among HOs for media attention and donor money, they tend to cooperate at the operational level. Co-location with other organiza-
tions in humanitarian clusters or hubs can provide benefits through complementary resources, such as shared technological resources and access to joint IT support, help with coordinating logistics and security resources, reduced operational costs, and facilitation of knowledge sharing. Copenhagen is one such example, which hosts both UNHCR and most other UN organizations, as part of the UN city established in 2013.

2.4.3 Model

In this section, we describe a two-stage mathematical programming model for solving a FLP. To deal with uncertainty of ER demand, we developed a scenario-based two-stage stochastic program with the aim of making robust first-stage decisions so that the second stage is feasible under all scenarios. According to Liberatore et al. (2013), such a robust programming approach is a common methodology for dealing with uncertainty when uncertain input distributions cannot be estimated reliably.
Sets

S  Set of suppliers
G  Set of potential global warehouses
A  Set of all demand points including ongoing and emergency
K  Set of emergency response scenarios

Parameters

\( f_g \)  Fixed cost of opening a potential global warehouses in USD
\( v_g \)  Variable cost of opening a potential global warehouses in USD
\( C_{ER}^{gj} \)  Transportation cost from global warehouse \( g \) to demand point \( j \) using express air shipment in USD
\( C_{ER}^{gj} \)  Transportation cost from global warehouse \( g \) to demand point \( j \) using express land shipment in USD
\( C_{so}^{gj} \)  Transportation cost from supplier \( s \) to demand point \( j \) using normal shipment in USD
\( C_{go}^{gj} \)  Transportation cost from supplier \( s \) to global warehouse \( g \) using normal shipment in USD
\( l^e_{gj} \)  Lead time from global warehouse \( g \) to demand point \( j \) using express air shipment in days
\( l^e_{gj} \)  Lead time from global warehouse \( g \) to demand point \( j \) using express land shipment in days
\( l^n_{gj} \)  Lead time from global warehouse \( g \) to demand point \( j \) using normal shipment in days
\( l^n_{sj} \)  Lead time from Supplier \( s \) to demand point \( j \) using normal shipment in days
\( d_{do}^{o} \)  Demand for ongoing operations at location \( j \) in USD
\( d_{do}^{ER,k} \)  Demand for emergency response at location \( j \) under scenario \( k \) in USD
\( lim \)  Equal to 430,000 USD, the dollar value of ten TEU’s
\( Q_g \)  A large number, or capacity of \( g \)’th global warehouse in TEU
### Decision Variables

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_g$</td>
<td>The required amount of inventory to be stocked in global warehouse $g$ in TEU</td>
</tr>
<tr>
<td>$z_{sg}$</td>
<td>Number of containers to send from supplier $s$ to global warehouse $g$ with normal shipment in TEU</td>
</tr>
<tr>
<td>$x_{so,k}$</td>
<td>Percentage of ongoing demand to send from supplier $s$ to demand point $j$ with normal shipment in the $k$'th scenario</td>
</tr>
<tr>
<td>$x_{gj}$</td>
<td>Percentage of ongoing demand to send from global warehouse $g$ to demand point $j$ with normal shipment in the $k$'th scenario</td>
</tr>
<tr>
<td>$x_{ER,k}^{gj}$</td>
<td>Percentage of emergency demand at location $j$ satisfied from global warehouse $g$ with express air shipment in the $k$'th scenario</td>
</tr>
<tr>
<td>$e_{x_{gj}}^{ER,k}$</td>
<td>The excess percentage of emergency demand at location $j$ satisfied from global warehouse $g$ with express land shipment (after sending 10 TEU’s by express air) at scenario $k$</td>
</tr>
<tr>
<td>$y_g$</td>
<td>1, if global warehouse $g$ is opened 0, otherwise</td>
</tr>
<tr>
<td>$b_{z_{sg}}$</td>
<td>1, if containers are sent from suppliers $s$ to global warehouse $g$ 0, otherwise</td>
</tr>
<tr>
<td>$b_{x_{so,k}}$</td>
<td>1, if containers are sent from supplier $s$ to demand point $j$ 0, otherwise</td>
</tr>
<tr>
<td>$b_{x_{gj}}$</td>
<td>1, if containers are sent from global warehouse $g$ to demand point $j$ 0, otherwise</td>
</tr>
<tr>
<td>$b_{x_{ER,k}^{gj}}$</td>
<td>1, if excess containers are sent by express air from global warehouse $g$ to demand point $j$ 0, otherwise</td>
</tr>
<tr>
<td>$b_{e_{x_{gj}}^{ER,k}}$</td>
<td>1, if excess containers are sent by express land from global warehouse $g$ to demand point $j$ 0, otherwise</td>
</tr>
</tbody>
</table>
2.4.3.1 Problem definition

UNHCR deals with two sets of supply points: global warehouses and supplier locations. There are also two types of demand points: i) OO, where demand is stable; and ii) ER, for which locations and magnitudes are uncertain when decisions are being made. OO demand can be met through supplier locations or inventory stocked at the global warehouses, whereas ER points are supplied only by the global warehouses. We treat the historic annual ER demand volumes as scenarios with equal probability of occurring (see Section 2.4.5 for details). Inventory at the global warehouses and in transportation is measured in twenty-foot equivalent units (TEU), a commonly used reference of volume in containerized shipping. We then convert the quantity of goods from TEU to the corresponding USD value based on estimates provided by UNHCR. Opening of each warehouse incurs a fixed cost and the inventory held at a location incurs a cost per unit stored in that warehouse. Shipping rates and lead times differ between transportation modes. Normal shipment uses surface transportation (by road or sea) to satisfy OO demand and is generally cheaper, but the lead time is longer. UNHCR regards road and sea transport as alternatives or complementary modes with minimal cost variance and does not distinguish between them in the operational context. Express shipment employs air or road transportation (if over a short distance) to satisfy ER demand, and it is associated with higher transport costs and shorter lead times. Owing to the higher costs, express shipment via air to any particular demand point is constrained by the availability of funding and cannot exceed 10 TEUs per disaster event. Additional TEUs are sent via surface. The goal of the proposed multi-objective model is to satisfy all ongoing and emergency demand in the fastest manner while incurring the minimal cost. To handle the dual objectives of lead time and cost, we developed three related two-stage mixed
integer mathematical programs. The first program solves the FLP to minimize the expected total cost, while disregarding lead time. The analysis in the second model is based on minimizing the expected lead time of the chosen supply chain network, while disregarding cost, and in the third, the same lead time objective is used, while constraining the supply chain budget based on the optimal minimum cost value obtained from the first program plus varying mark ups. In the rest of this section we briefly describe these three models. Computational results based on UNHCR data are discussed in Section 2.4.5. An IBM ILOG CPLEX machine with 8 GB of RAM was used for solving each of the three mathematical models with maximum of 19,288 binary, integer, and continuous variables and 10,968 constraints. All model instances were solved within seconds.

2.4.4 Minimum Expected Total Cost Model

Our formulation in 2.1 and 2.2 comprises of three sets of variables. First, for every node in set G, we define a binary variable $y_g$, where, $y_g = 1$ if the $g^{th}$ global warehouse is opened. Next, nonnegative integer flows on arcs from supplier locations to global warehouses determine inventory levels at the opened warehouses. Last, we define nonnegative continuous variables over each arc going from suppliers and warehouses to demand points. These variables represent the required percentage of OO or ER demand sent through each of these arcs. We assume no capacity restrictions at the suppliers, warehouses, or transport arcs. Below we show the minimum
expected total cost formulation.

\[
\min \sum_{g \in G} f_g y_g + \sum_{g \in G} \sum_{s \in S} C_{sg} z_{sg} + \sum_{g \in G} v_g I_g + \\
\sum_{k \in K} \sum_{j \in A} \alpha_k \left( \sum_{g \in G} \left( C_{gj}^o x_{gj}^o + \sum_{s \in S} C_{s, gj}^o x_{gj}^o \right) d_j^o + \\
\sum_{g \in G} \left( (C_{gj}^{ER} x_{gj}^{ER})((1 - \gamma_j^k)\text{lim} + d_j^{ER,k} \gamma_j^k) + (C L_{gj}^{ER} e x_{gj}^{ER,k})(1 - \gamma_j^k)(d_j^{ER,k} - \text{lim}) \right) \right)
\]

(2.1)

Subject to:

\[
\sum_{g \in G} (x_{gj}^o d_j^o) + \sum_{s \in S} (x_{sj}^o d_j^o) = d_j^o, \forall j \in A,
\]

(2.2a)

\[
\sum_{g \in G} (x_{gj}^{ER,k} d_j^{ER,k} + e x_{gj}^{ER,k} d_j^{ER,k}) = d_j^{ER,k}, \forall k \in K, j \in A,
\]

(2.2b)

\[
\sum_{g \in G} (x_{gj}^{ER,k} d_j^{ER,k}) = (1 - \gamma_j^k)\text{lim} + d_j^{ER,k} \gamma_j^k, \forall j \in A, k \in K,
\]

(2.2c)

\[
\sum_{g \in G} (e x_{gj}^{ER,k} d_j^{ER,k}) = (1 - \gamma_j^k)(d_j^{ER,k} - \text{lim}), \forall j \in A, k \in K,
\]

(2.2d)

\[
I_g \leq Q_g y_g, \forall g \in G,
\]

(2.2e)

\[
\sum_{s \in S} z_{sg} = I_g, \forall g \in G,
\]

(2.2f)

\[
\sum_{j \in A} (x_{gj}^o d_j^o + x_{gj}^{ER,k} d_j^{ER,k} + e x_{gj}^{ER,k} d_j^{ER,k}) \leq 43000 I_g, \forall k \in K, g \in G,
\]

(2.2g)

\[
x_{s, gj}^o, x_{gj}^o, x_{gj}^{ER,k}, e x_{gj}^{ER,k} \in [0, 1], \forall k \in K, s \in S, g \in G,
\]

(2.2h)

\[
z_g, I_g \text{ integer, } \forall g \in G,
\]

(2.2i)

\[
y_g \text{ binary, } \forall g \in G,
\]

(2.2j)
The objective function \((2.1)\) in TC* minimizes the total cost associated with opening warehouses, holding necessary inventory at opened warehouses, shipping cost from suppliers to global warehouses, expected shipping cost from suppliers and global warehouses to OO points, and expected shipping cost from global warehouses to ER points. This objective is subject to several constraints. Constraints \((2.2a)\) and \((2.2b)\) require satisfying all OO and ER demand. Constraints \((2.2c)\) and \((2.2d)\) ensure that if ER demand is greater than 10 TEUs, 10 TEUs of the total demand must be sent by express air shipment and the rest by express road shipment. If ER demand is less than 10 TEUs, the entire demand is sent by air. Constraint \((2.2e)\) allows only opened warehouses to have positive inventory. Constraint \((2.2f)\) requires the inventory stocked at any global warehouse to be equal to the sum of shipments from suppliers to said warehouse. Constraint \((2.2g)\) requires that the total OO and ER shipments from any given global warehouse be less than or equal to its inventory. Constraint \((2.2h)\) denotes \(x_{oo, k}^{oo,k}, x_{oo, k}^{oo, k}, x_{oo, g}^{ER, k}, x_{oo, g}^{ER, k}\) as nonnegative continuous \(\text{Cow variables, (2.2i) denotes } I_g \text{ and } z_{sg} \text{ as nonnegative integer variables, and (2.2j) denotes } y_g \text{ as binary variables.}

\[2.4.4.1 \text{ Minimum Expected Lead Time Model}\]

The formulation LT* minimizes only the total lead time associated with all used arcs from a supply point to an OO or ER demand point (Fig.B3). The objective function given in \((F2)\) minimizes the expected lead time of shipping from suppliers to OO demand points, from global warehouses to OO demand points, and from global warehouses to ER points. We do not consider the lead times of shipments from supplier locations to global warehouses because we assume that replenishment of global warehouses is an ongoing process that does not impact supply chain responsiveness. This
assumption is based on the fact that UNHCR (i) keeps a buffer stock for 600,000 beneficiaries in its global warehouses and (ii) conducts forward planning and preemptive timeslot-based order placement, whereby country operations do not experience any increase in lead time for deliveries from global warehouses. Analysis of the inventory ordering policies of the global warehouses are beyond the scope of this paper because we only focus on high-level planning decisions and, therefore, consider only the total annual inventory levels at the global warehouses.

\[
\min \sum_{k \in K} \sum_{j \in A} \alpha_k \left( \sum_{g \in G} p_{gj}^k b_{x_{gj}}^{oo,k} + \sum_{s \in S} l_{s}^k b_{x_{sj}}^{oo,k} + \sum_{g \in G} t_{gj}^k b_{x^{ER,k}_{gj}} + \sum_{g \in G} L^k_{gj} b_{cx^{ER,k}_{gj}} \right)
\]  

(2.3)

Subject to:

\[b_{x_{gj}}^{oo,k} \geq x_{gj}^{oo,k}, \forall g \in G, j \in A, \]  

(2.4a)

\[b_{x_{sj}}^{oo,k} \geq x_{sj}^{oo,k}, \forall s \in S, j \in A, \]  

(2.4b)

\[b_{x^{ER,k}_{gj}} \geq x^{ER,k}_{gj}, \forall g \in G, j \in A, \]  

(2.4c)

\[b_{cx^{ER,k}_{gj}} \geq cx^{ER,k}_{gj}, \forall g \in G, j \in A, \]  

(2.4d)

\[b_{x_{gj}}^{oo,k}, b_{x_{sj}}^{oo,k}, b_{x^{ER,k}_{gj}}, b_{cx^{ER,k}_{gj}} \text{ binary}, \forall g \in G, \]  

(2.4e)

2.4.4.2 Budget Constrained Minimum Expected Lead Time Model

Dual objectives of lead time and cost are handled by a third model, in which the same lead time objective function as in LT* is minimized subject to all previous constraints and an additional budget constraint (Fig. B4). This constraint ensures that the total cost incurred under this model is less than or equal to the optimal value of TC* plus a percentage, b. In our computational experiments for analyzing
the trade-off between cost and lead time, we varied $b$ such that the right-hand side of constraint (16) ranged from the optimal objective value of $TC^*$ (minimum possible cost) to the cost incurred in the optimal $LT^*$ solution. Thus, we approximated the Pareto frontier with a set of efficient solutions, obtained by solving the model with $b$ values ranging from 0 to 0.2 with increments of 0.0005, such that no other solution with better cost as well as better lead time exists.

$$\sum_{g \in G} f_g y_g + \sum_{g \in G} \sum_{s \in S} C_{sg} z_{sg} + \sum_{g \in G} v_g I_g +$$

$$\sum_{k \in K} \sum_{j \in A} \alpha_k \left( \left( \sum_{g \in G} C_{oo, g,j} x_{oo, g,j} + \sum_{s \in S} C_{oo, s,j} x_{oo, s,j} \right) q_{oo} + \right.$$

$$\left. \sum_{g \in G} \left( (C_{g,j} e_{g,j}^{ER,k})(1 - \gamma_{j}) \lim(1 - \gamma_{j}) \right) (d_{j}^{ER,k} - \lim) \right) \leq TC^*(1 + \beta)$$

2.4.5 Case for model testing

In this section, we present the case on which the model was tested; it is based on real data collected from or estimated in agreement with UNHCR (Appendix B). The main aspect of the analysis is to compare UNHCR’s existing warehouse network constituting 7 global warehouses with a redesigned network containing four new locations (Figure 2-2). The alternative locations were agreed upon in discussion with UNHCR: Subang (Malaysia) means possible co-location with UNHRD (www.unhrd.org), Algeciras (Spain) is good for transshipment (UNHRD may decide to relocate their warehouse from Las Palmas for the same reason), while Karachi (Pakistan) and Djibouti (Djibouti) both are big distribution points for OO. UNHCR has practically no operations in the Americas, and there are no plans for a global warehouse in this region.
The resulting model includes the 11 candidate warehouse locations, 56 demand points for OO, 76 demand points for ER, and 14 supplier points. OO demand was calculated as a three-month average (in USD) per distribution point (one per country) based on historic data from the years 2011-2013. We used this data to create three emergency demand scenarios represented as three-month averages per distribution point (Figure 2-3), yielding representative variability and trends. Each year constitutes one scenario with an occurrence probability of 0.3333.

Data on transportation and warehousing were collected and used to calculate/estimate costs, capacities, and lead times (Appendix A). In agreement with UNHCR, it was decided to quantify four main influencing factors by estimating location-specific percentages and adding/subtracting from variable or fixed costs (Table 2.4). We see that the factors have different weights for the same location, which implies that it is important to incorporate more than one factor. In particular, hardship impacted a
greater number of locations than security, and locations with high security cost such as Karachi had relatively lower accessibility cost than other locations, for example, Douala and Isaka.

The first, differentiated access to human resources, was quantified using UN’s hardship classification (policy duty stations are classified from A to E, with the latter being the hardest) and payment schemes (ICSC, 2013) to calculate the additional percentage of staff’s share of fixed cost and variable cost. The qualitative study suggested pilferage as an important factor, and we undertook an analysis of UNHCR insurance claim history over 2011–2014 to establish potential patterns varying with location. No such pattern was identified, and in agreement with UNHCR, we in-
cluded pilferage in the differentiated security cost to account for the extra security measures taken to avoid loss (e.g. fences and guards). Security cost was quantified using the UN Security Management System Security Policy Manual with levels ranging from 1 to 6, with 6 representing the most dangerous environment (United Nations Department of Safety and Security, 2011) together with the most updated security rankings (trip.dss.un.org) for the locations analyzed, assuming in agreement with UNHCR, a non-linear relationship with no extra cost for levels up to 3, 15% for level 4, 45% for level 5, and a “no-go” notification for level 6 (i.e., no global warehouse in such locations). For accessibility, we used country scores from the Logistics Performance Index (lpi.worldbank.org) to establish the percentages of additional variable warehousing cost. Finally, in agreement with UNHCR, co-location was accounted for by assuming 50/50 split of fixed warehousing cost with the other organization.

2.5 Computational Results

This section presents the three key insights derived from the computational results relating to the objectives presented in Section 2.2a: i) Quantification of the impact of an expanded network on lead time and cost; ii) advantage of joint prepositioning; and iii) impact of the factors on warehouse locations, that is, network configuration.

Key insight 1: Network expansion reduces costs and shortens lead times

Figure 2-4 shows the trajectories and efficient frontiers of the two different network configurations when \( \beta \) is varied from 0 to 0.2. The rightmost curve represents the existing network with seven candidate warehouses, and the leftmost curve represents the expanded network with an additional four warehouses. Each point on the frontiers shows the optimal lead time for the third model associated with a different
\( \beta \) under the budget constraint that allows a cost of \((1+\beta)\) times the optimal value obtained using the TC* model. That is, the minimum lead time is given on the y axis when spending is constrained by cost, which is represented on the x axis. We have highlighted one point \((\beta = 0.015)\) which we will use when discussing further key insights because it represents a good trade-off between cost and lead time.

The results indicate that when optimizing cost and lead time, expanding the existing network by adding the four candidate locations enables a 31% cost reduction and an 18% lead time reduction when \(\beta = 0.015\). When optimizing LT*, the expanded network leads to a 28% reduction in lead time. As expected, there is a trade-off between lead time and cost for both the existing and the expanded networks. It is interesting to note, however, that the increase in cost is relatively small compared to the savings in lead time. Relaxing the budget constraint by just a few percent and allowing for an expanded network of global warehouses makes it possible to achieve substantial savings in terms of lead time. In other words, based on the computational results, we see that UNHCR, as well as other HOs using prepositioning, should consider establishing additional global warehouses closer to points
with high demand. The computational experiments indicate that UNHCR should consider closing a few of the warehouses in the existing network and open new ones (Table 2.5). For example, when optimizing TC* ($\beta = 0$) for the expanded network, Dubai and Douala were replaced by Algeciras, Subang, and Djibouti. Copenhagen helped improving the lead time, but it increased costs; therefore, it was not viable without a further increase in budget (at $\beta = 0.035$).

A few of the open/close decisions change more than once as the budget increases. It is beneficial to open Dubai when $\beta$ is large enough ($\beta = 0.015$). However, when the budget increases further and is sufficient to support Karachi, Dubai is no longer needed because Karachi is better from the lead time perspective. Having both Karachi and Dubai open at the same time is not feasible budget-wise. The above results also indicate that network configuration is sensitive to the available budget.

**Key insight 2: Joint prepositioning for ER and OO demand reduces lead times and total costs**

The second key insight from the model relates to the benefits of joint prepositioning for satisfying OO and ER demand. By joining the OO and ER supply chains, prepositioned stock can be used for both long- and short-term operations. To exemplify, we show what happens at $\beta = 0.015$ in Table 2.6, where the second column indicates whether a warehouse location is opened or closed. The third column indicates the number of TEUs shipped from each location to satisfy the total OO and ER demands, while the fourth and fifth columns show the amounts of OO and ER in USD handled through each warehouse, that is, the sum of the demands in USD provisioned from each warehouse to all ER/OO demand points in all scenarios.
We see that eight of the warehouses satisfy OO demand and seven satisfy both OO and ER demands. In our computations, although the source of satisfying OO demand can be different, for $\beta = 0.015$, the OO demand points are served from the same global warehouses across scenarios. In comparison, the global warehouses fulfilling ER demand vary significantly with scenarios, primarily because ER demand is very volatile across the employed scenarios (Figure 2-3). For example, there was almost no ER demand in Jordan in 2011, whereas in 2012 and 2013, the average three-month ER demand corresponded to USD 5.17 million and USD 13.77 m, respectively. From the computational evidence, we conclude that by introducing stable OO demand in the network, we can justify opening additional warehouses, which, in turn, leads to shorter lead times and lower total cost for both OO and ER (Figure 2-4).

Table 2.5: configurations under different margins within the budget constraints (1 = location is opened; 0 = location is closed)

<table>
<thead>
<tr>
<th>Location included in modeling</th>
<th>$\beta = 0$</th>
<th>$\beta = 0.001$</th>
<th>$\beta = 0.015$</th>
<th>$\beta = 0.025$</th>
<th>$\beta = 0.035$</th>
<th>$\beta = 0.045$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accra (Ghana)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Algeciras (Spain)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Amman (Jordan)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Copenhagen (Denmark)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Djbouti (Djibouti)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Douala (Cameroon)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Dubai (UAE)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bukit (Tanzania)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Karachi (Pakistan)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Nairobi (Kenya)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sabang (Malaysia)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total open warehouses</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

**Key insight 3: Contextual factors matter**

The third insight relates to the importance of contextual factors, in particular, political and security factors. Table 5 summarizes the computational results for $\beta = 0.015$, where each column represents a separate run of the model. The second column considers all four factors. Thereafter, we exclude only hardship, then only
security, then only co-location, then only accessibility, and finally, we consider no factor (i.e., we set the factors to 0% benefit or extra cost). The results show that all factors influence network configuration. For example, Karachi is opened, while Douala and Dubai are closed in case either hardship or security is not accounted for. Furthermore, not accounting for the co-location factor leads to opening of Nairobi and closure of both Isaka and Dubai. Notably, Accra and Subang are kept open in spite of a 100% increase in fixed warehousing cost (because of the absence of co-location benefits), implying that they represent key locations in the network to lower costs and lead times. The exclusion of accessibility seems to have less of an effect on location selection. Finally, the exclusion of all factors implies that both Dubai and Douala are closed in contrast to the case when all factors are accounted for. It should be noted that the effect on lead time varies depending on which factor is excluded from the model. Particularly, the results show that lead time is reduced by 7.7% when the extra cost of hardship is not considered. By excluding hardship, the total cost decreases and the extra budget available can be used to open Karachi, Pakistan, which is closer to several demand points, thus helping reduce lead time. In contrast, excluding co-location increases the total cost as well as the lead time (4.7%).
To check whether the choice of \( \beta \) changes the impacts of the factors on network configuration, that is, to test whether our computational results related to factor analysis are robust under different budget levels, we ran the same computations with different \( \beta \) values. We found that all factors impact warehouse decisions at all budget levels, with the exception of settings with very low \( \beta \) values. This is because if the budget is very low, only a very few feasible solutions exist to begin with; for example, with \( \beta = 0.001 \), no network that can satisfy all demand if co-location is excluded. Further, the way these factors impact network configuration might change with higher \( \beta \) values because higher budgets increase financial flexibility and make the decision to open/close a location less dependent on the inclusion/exclusion of extra costs, for example, those related to hardship.

2.6 Conclusions, Implications, and Future research

We introduced an optimization model for quantifying the impact of an expanded network and determining the best locations for joint prepositioning of relief items to serve both short- and long-term operations. The model was developed based on empirical evidence, and it offered data-driven insights on factors that significantly impact location choice in the humanitarian context. The results suggest that joint prepositioning allows the organization to open a greater number of global warehouses, while reducing costs and response times. Moreover, we found that factors related to security, accessibility, co-location, and human resources, when quantified and modeled, change the network configuration. This study contributed to research on the topic at hand in two ways. First, we developed a framework and a data-driven model that integrate the long-term and the emergency relief supply chains, accounting for
both response time and cost. Second, our study helps fill the gap in humanitarian network design literature by including factors that influence warehouse locations for prepositioned stock in the decision models. While our analysis is based on data from a single case study, the framework and model, as well as the process through which the factors are quantified, can be generalized and used by other HOs aiming to reduce costs, while maintaining speedy response in both short- and long-term operations. Many of these organizations such as WFP and UNICEF have similar organizational structures and HR policies, and they operate in the same countries under comparable political and security conditions as UNHCR. The main practical contribution of this study is to provide UNHCR and other organizations with decision support in network redesign, accounting for two matters of (increasing) practical importance: improving performance through joint prepositioning and quantification of factors that should be accounted for when deciding warehouse locations. Joint prepositioning of stock can, in addition to reducing cost and lead time, enable reduction of country warehouses and inventory of consigned stock. A greater amount of un-consigned stock improves flexibility because it can be redirected toward ongoing operations in the region and/or emergencies, rather than awaiting an emergency that might not occur. Another practical implication is that stock in strategic locations can significantly cut lead times in ongoing operations. This is because budget often is made available at a late stage, with little time left to fulfill demand. Meanwhile, lead times from suppliers can be several weeks, whereas a nearby global warehouse would cut this down to a few days. The perceived response times for operations could thus be reduced by pre-ordering based on historical demand. Thereby, country operations may benefit from “guaranteed” availability within a short time frame, which reduces the perceived need for large local stocks and reduces the negative effects of delays in funding. Combining the two flows may also allow for economy of scale and better
utilization of production capacity at suppliers’ sites through improved planning and advance placement of non-rush orders. Finally, by allowing more time for planning and locating global warehouses closer to emergency demand points, we can increase the use of cheaper transport and reduce the use of expensive air transport.

Our model can be enhanced in several ways. One important aspect is further validation of input data such as costs of human resources, normal versus express shipment, and setting up and operating global warehouses, as well as further testing and validation of the quantification of factors. Moreover, it would be valuable to extend the model to include country warehouses, thus allowing for an analysis of the benefits of reducing consigned stock at this level and instead increasing the amount of un-consigned stock in global warehouses. Further, the model could be tested by using data from other HOs such as WFP, UNICEF, and IFRC. It could also be extended for supporting operational decision making in addition to strategic planning, which it currently supports. Finally, it is critical to consider scenario planning and development for predicting future ER demand. Scenarios could be developed for example by combining statistics provided by the Internal Displacement Monitoring Centre (IDMC) with contingency plans for specific countries and vulnerability indexes.
Chapter 3

Forecasting Refugees Emergency Relief Response at United Nations High Commissioner for Refugees (UNHCR)

3.1 Introduction and Motivation

Humanitarian emergencies force larger number of people to flee their homes every year. As a result, Emergency Relief (ER) operations are increasingly taking up a greater share of Humanitarian Organizations’ (HO) budgets. Similar to other HOs, United Nations High Commissioner for Refugees (UNHCR) divides its resources between emergency response and long-term ongoing operations. ER operations, as the time-sensitive component of UNHCR, deal with sudden-onset disasters due to persecution, conflict, violence, or human rights violations where fast humanitarian
assistance is essential. On the other hand, Ongoing Operations (OO), as the cost-sensitive component of UNHCR, support long-term efforts such as camp operations. As the number of refugees affected by humanitarian emergencies has grown drastically, particularly after 2005 [32], UNHCR has introduced new initiatives focusing on improving emergency response planning and preparedness [31].

The goal of ER planning and preparedness initiatives are not only to shorten the response time, but also to improve the percentage refugees receiving aid. The UN Secretary General and United Nations Office for the Coordination of Humanitarian Affairs (OCHA) reported that $15.2 billion was spent on humanitarian emergencies in 2018 [35,28]. Although the number of people in need was reported as 135.3 million, only 97.9 million people [35] (%73) received aid. In this report, UN Secretary-General suggested a collection of efforts in ER planning and preparedness to improve the efficiency of humanitarian emergency assistance. Among suggested efforts were: 1) shifting from reaction to anticipation, and 2) improving humanitarian coordination and response. So far, the majority of the research on emergency response planning and preparedness and overall humanitarian response have been focused on improving coordination and response rather than the anticipation aspect. Most of these efforts optimize coordination through prescriptive models that take the forthcoming emergencies as predetermined parameters in the models.

The effectiveness of policies resulting from such prescriptive models is heavily dependent on the accuracy of predictions for future ER response. Surprisingly, despite the large spike in popularity of ER planning and preparedness research in recent years, the problem of forecasting emergency response for refugees is understudied [94]. As an example of this trend, optimization models for emergency planning and
preparedness suggest that pre-positioning relief items is one of the most effective actions for fast, responsive, and effective delivery of aid. Having studied the performance of CARE, Duran et al. (2011) showed the importance of pre-positioning the relief items on shortening the average emergency response time to people affected by natural disasters [27]. In a recent study, Jahre et al. (2016) also showed the significance of joint pre-positioning of relief items for both emergency and ongoing operations on UNHCR’s total cost and response time for refugees affected by conflicts [47]. Various studies have investigated the effects of pre-positioning of supplies [23], warehouse locations [95], resource allocations [31], transportation planning [22], partnerships developments [18], and integration of multiple supply chain resources and redesign of network configurations [47]. Nevertheless, none of these works has focused on the critical issue of forecasting ER response, investigating factors that play a key role in development or identification of future emergencies, nor predicting the risks associated with countries experiencing emergencies.

Moreover, among academic studies and practices in forecasting ER response, natural emergencies, as opposed to emergencies caused by conflicts and human rights violations, have received more attention and continues to attract more funds, donations, media coverage [82], and even dedicated software [62, 18]. However, persecution, conflict, violence, or human rights violations are the main drivers of ER response at HOs such as UNHCR. Even in the case of natural disasters, communities in conflict–ridden areas are more vulnerable to natural hazards.

In this paper, we explore the relationship between countries’ economical, societal, political, geographical, environmental, infrastructural, employment, human rights, and refugee movement situations with their emergency response. Using this rela-
tionship, we predict the likelihood of experiencing humanitarian emergencies that result in ER response at each country in the following year, estimate the value of ER response, and discover identifiers of emergencies. We describe the associations between ER response and its main predictors which lead to valuable insights. Our findings emphasize the key role of factors introduced in this study, many of which are not used by humanitarian organizations.

This study focuses on forecasting emergency response for UNHCR, which covers 74.8 million people worldwide and has more than 16,803 personnel across 134 countries in 2018. Given the importance of resource allocation in ER operations, accurate prediction of ER response can serve millions of people more swiftly and efficiently.

3.2 Literature Review and Our Contribution

Our adopted strategy for predicting future emergency response is inspired by semi-continuous data literature and economics. In this area, the data that takes continuous and positive values but has a substantial proportion of values at zero is often called “zero-inflated” [64]. Zero-inflated data causes challenges in predictive modeling such as producing biased parameter estimates and standard errors. In the case of emergency response, while many countries have ER response, many others (or often the majority) do not require ER assistance at all. This creates a clump at zero in the frequency distribution of ER data. The literature of zero-inflated variables is divided into two groups differing on the assumption that zero and positive values are drawn from the same or separate distributions. Shono in 2008 applied and compared both approaches to predict catch per unit effort, which is a zero-inflated variable representing
the stock abundance \cite{31}. In his analysis, the approach with the same distribution (Tweedie regression model) generates better performance in terms of squared errors compared to the classic implementation of the approach with separate distributions (Bernoulli-lognormal two-part regression model).

On the empirical side, papers that considered uncertainty and volatility of ER response in field are as follows: Dalal et al. (2017) embedded multiple scenarios into humanitarian logistics models using stochastic modeling, robust modeling, or weighted sum of the two frameworks\cite{23}. Zhang (2007) and Sheu (2010) developed models for dynamic demand forecasting through times series methods such as Auto-Regressive and Integrated Moving Average, exponential smoothing models, and independent identically distribution models \cite{93, 80}. Sheu (2010) and Jiang (2009) used Fuzzy theory-based methods to estimate demand as a time-varying variable based on accumulated numbers of fatalities \cite{80, 49}. Nadi and Edrisi (2017), Taskin et al. (2011), Mishra and Desai (2006), Yang et al. (2016) identified an effective way of risk analysis and assessment for natural disasters using Markov decision process, Bayesian networks, and neural networks \cite{66, 84, 65, 92}.

The analysis of this paper improves strategies for forecasting refugees emergency response. We go deeper into shifting from reaction to anticipation aspect of emergency response planning and preparedness. Relative to these literature, our study makes three main contributions. First, to our knowledge, none of the prior studies on refugees emergency response forecasting have been conducted at the scale of this study, characterizing country-year level data from multiple sources and humanitarian organizations with the focus on humanitarian emergencies. Second, we show that none of the standard forecasting techniques in statistical modeling and machine learning can successfully predict the future ER response across different countries.
Therefore, we created a new framework which utilized the existing methods and also discussed the necessity of adopting this framework in emergency response forecasting. Finally, we identified the best set and format of data to use for I predicting the risks associated with future refugee–related emergencies, II forecasting the value of required ER response, and III identifying the countries needing ER assistance. We find that different set of data is better suited for each of these three purposes.

3.3 Data

3.3.1 Publicly Available Data

Our data set is comprised of several HOs’ databases. We used publicly available data collected by UNHCR [33], World Bank (WB) [8], Uppsala Conflict Data Program (UCDP) [86], and Index for Risk Management (INFORM) [34]. Different HOs have adopted varying definitions for terms and indices and sometimes data formats can change significantly from one to the other. Consequently, data collected from different sources may contain inconsistent or overlapping information. We used manuscripts, yearly reports, published papers, and websites to choose our data from sources that are compatible in their definitions throughout the years. We produced a raw data set with each record explaining a wide spectrum of a country’s characteristics including their economical, societal, political, geographical, environmental, and infrastructural situations each year. We also included each country’s refugee movement and employment and human rights conditions. In addition, to collect budget related data, we collaborated with UNHCR, who provided us with invaluable data on amounts spent on ongoing operations and emergency response operations.
These characteristics are collectively referred to as “features” or “explanatory variable” interchangeably throughout the paper. Each record in the data set represents a country in a particular year. Such data format is commonly known as a “country-year level” data. Tables 3.1, 3.2 and 3.3 outlines the features employed in the raw data set, including their definition, summary statistics, and the source of each feature.

As mentioned above, different sources contributed to data set in various aspects. For example, Index for Risk Management (INFORM) provides features representing the environmental, human rights, political, socio-economical, institutional, and infrastructural aspects. INFORM, a collaborative project between the Inter-Agency Standing Committee (IASC) and the European Commission, is a global crisis risk index widely utilized in humanitarian crises and disaster management. INFORM model is based on risk concepts derived from three dimensions of risk: hazards and exposure, vulnerability, and lack of coping capacity. Figure 3-1 represents different dimensions, categories, and components of INFORM Index. Each country is given an INFORM risk grade from 0-10 with 0 being low risk and 10 being extremely high risk for any natural disasters or humanitarian crises. The data provided by INFORM includes 10 features developed for 191 countries since 2013. In our raw data set, we include INFORM risk index and its dimensions.

<table>
<thead>
<tr>
<th>Variable Description</th>
<th>Source</th>
<th>Source Type</th>
<th>Range</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amounts spent on ongoing operations in the current year</td>
<td>UNHCR</td>
<td>Continuous</td>
<td>4.01E+07, 9.69E+06</td>
<td>3.98E+06</td>
<td>0.00E+00</td>
<td>4.01E+07</td>
<td></td>
</tr>
<tr>
<td>Amounts spent on emergency relief operations in the current year</td>
<td>UNHCR</td>
<td>Continuous</td>
<td>6.79E+07, 8.69E+05</td>
<td>3.88E+06</td>
<td>0.00E+00</td>
<td>6.79E+07</td>
<td></td>
</tr>
<tr>
<td>Number of Refugees in UNHCR’s persons of concern</td>
<td>UNHCR</td>
<td>Continuous</td>
<td>2.50E+06, 9.35E+04</td>
<td>2.64E+05</td>
<td>0.00E+00</td>
<td>2.50E+06</td>
<td></td>
</tr>
<tr>
<td>Number of Returning Refugees in UNHCR’s persons of concern</td>
<td>UNHCR</td>
<td>Continuous</td>
<td>7.63E+06, 2.69E+05</td>
<td>9.16E+05</td>
<td>0.00E+00</td>
<td>7.63E+06</td>
<td></td>
</tr>
<tr>
<td>Number of Internally displaced persons in UNHCR’s persons of concern</td>
<td>UNHCR</td>
<td>Continuous</td>
<td>8.23E+05, 2.19E+04</td>
<td>3.38E+04</td>
<td>0.00E+00</td>
<td>8.23E+05</td>
<td></td>
</tr>
<tr>
<td>Number of Asylum-seekers in UNHCR’s persons of concern</td>
<td>UNHCR</td>
<td>Continuous</td>
<td>1.10E+06, 1.29E+04</td>
<td>6.84E+03</td>
<td>0.00E+00</td>
<td>1.10E+06</td>
<td></td>
</tr>
<tr>
<td>Number of Stateless persons in UNHCR’s persons of concern</td>
<td>UNHCR</td>
<td>Continuous</td>
<td>9.38E+05, 2.98E+04</td>
<td>1.26E+05</td>
<td>0.00E+00</td>
<td>9.38E+05</td>
<td></td>
</tr>
<tr>
<td>Number of Others of concern in UNHCR’s persons of concern</td>
<td>UNHCR</td>
<td>Continuous</td>
<td>9.38E+05, 2.98E+04</td>
<td>1.26E+05</td>
<td>0.00E+00</td>
<td>9.38E+05</td>
<td></td>
</tr>
<tr>
<td>Amounts spent on emergency relief operations in the following year</td>
<td>UNHCR</td>
<td>Continuous</td>
<td>6.79E+07, 9.20E+05</td>
<td>3.94E+06</td>
<td>0.00E+00</td>
<td>6.79E+07</td>
<td></td>
</tr>
<tr>
<td>Number of individuals assisted by UNHCR by the end of current year</td>
<td>UNHCR</td>
<td>Continuous</td>
<td>2.50E+06, 9.35E+04</td>
<td>2.64E+05</td>
<td>0.00E+00</td>
<td>2.50E+06</td>
<td></td>
</tr>
</tbody>
</table>

Uppsala Conflict Data Program (UCDP) supplied us with a feature regarding
organized violence and civil war. UCPD has recorded ongoing violent conflicts since the 1970’s. We use data on the the count of battle related deaths (BRD) (> 25) in 191 countries, recorded since 1989. The UCDP classifies a conflict as active for both state-based and non-state-based violence if there are at least 25 battle related deaths per calendar year in one of the conflict locations.

The World Bank (WB) provided us 15,576 unique records concerning economical, societal, and employment conditions. These features are gathered from 264 countries
for years 1960 to 2018. For the purpose of this study, we selected the following eight features from the WB database: GDP, GDP Growth, GNI, Inflation, Labor Force, Population, and Unemployment.

Finally, publicly available UNHCR data covers features describing the number of forcibly displaced people and refugee movements. UNHCR Population Statistics documents the locations of origin and residence (country or territory of asylum) for its population of concern along with the number of people who received assistance from UNHCR since 1951. This data set includes approximately 199 residence and 223 origin locations per year. Note that when merging data sets, the residence country and year is used as the key information to match UNHCR data with that of other sources. UNHCR categorizes its population of concern into seven groups and records the number of people under each category in residence countries throughout the years. These categories are: refugees, asylum seekers, internally displaced persons (IDPs), returned refugees, returned IDPs, stateless persons, and others of concern. A detailed and official explanation for people under each category can be found at popstats.unhcr.org/en/overview. In summary, each record of our raw data set includes a country name, the number people under each category and their origin countries, and a count of those who have received assistance through UNHCR by the end of each year.

Table 3.3: Feature Descriptions and Summary of Data Before Normalization (Continued)
3.3.2 Data Collected Through Collaboration with UNHCR

From our collaboration with UNHCR, we obtained data representing the amount of funds spent on OO and ER operations for the between 2011 and 2016. The UNHCR expenditure data on OO and ER of each year are added to the features discussed in 3.3.1. The ER data is of particular importance since our analysis is aiming at forecasting the future ER response in each country.

Our goal is to explain the relationship between countries’ features each year \(X_t\) and their ER response in the following year \(ER_{t+1}\). This relationship can be written as:

\[
ER_{t+1} = f(X_t) + \epsilon, \tag{3.1}
\]

in which \(f : X_t \rightarrow ER_{t+1}\) is a mapping of countries’ features in a year to ER of the following year. Since the expenditure data is used as the dependent variable in equation 3.1, the range of available data for the ER also restricts the range of useful records in the feature set. ER data is available for 94 countries between years 2011 and 2016. Consequently, having removed the countries and years in which the ER data is not available reduces the number of records to 470 rows. We use features of years 2011 to 2014 and ER response of 2012 to 2015 to train our models, and we test our models using features of 2015 to predict ER response of 2016. We also normalize continuous feature in training data set and record the parameters to normalize test data accordingly.

3.3.3 Extracting Refugee-Specific Information

Refugees-specific information can be extracted by explaining the dynamics of ER operations on the ground. Upon the occurrence of a conflict, internal displacement
of refugees can result in ER response within the country, while migration changes the location of ER response. As a result, the country receiving ER assistance may be different from the country of crisis. One of such examples is the crisis in Syria and the ensuing ER response in the residence countries of Turkey and Greece.

We extracted refugee-specific features from the existing information in the data set. These features are referred to as the engineered factors (ENG). Tables 3.1, 3.2, 3.3 specifies the list of engineered factors by ENG in the source column. The choice of such engineered factors is flexible and based on the best performance for a given problem. We identified six Engineered factors that improve model performance the most.

Also shown in Table 3.3, for each residence country in the data, we have the origin countries that refugees emigrate from as a set of binary features, denoted as “ORG_ (name of the country)”. To assist the model with the effect of immigration, we have introduced new variables to indicate the potential amount of need that can be transferred to each resident country by refugees migration. For example, “Source ER” is one of such features.

\[
Source\_ER_{j,t} = \sum_{i \in ORG\_Countries \ of \ (j,t)} ER_{i,t} \quad \forall j \in RES\_Countries, t \in Years
\]

(3.2)

For each residence country \( j \), the \( Source\_ER_{j,t} \) is defined as the total amount of ER response in its origin countries \( i \) in year \( t \). For example, Source ER at Australia in year 2011 equals to the sum of all ER response in the countries whose refugees immigrated to Australia in 2011. We add this engineered factor as the potential amount
of ER that can be transferred to country $j$ in year $t+1$ from all of its origin countries $i$ in year $t$. Similarly, for each residence country $j$, “Source $OO_{j,t}$” is defined as the total amount of OO response in its origin countries $i$ in year $t$. It represents the total amount of the OO response that can be transferred to country $j$ in year $t+1$ from all of its origin countries $i$ in year $t$.

Many of the residence countries have very similar set of origin countries, leading to multicollinearity. We added the refugees’ origin regions as a new group of binary engineered features (“ORG_(name of the region)”). This group of features, represent the geographical region of origin countries. The number of geographical regions is significantly smaller than the number of individual countries that refugees come from. This transformation reduces multicollinearity created by the binary origin country features. This is an attempt to examine whether substituting countries with regions in the data set is a good strategy.

We also created a new variable, “Number_of_Origin_Countries” ($NB_{OC}$), that simply represents the total number of countries sending refugees to a given residence country. This factor is included only if the regions-based approach is used rather than the origin country features. As when working with regions the binary features representing the origin countries is eliminated, it would be helpful to represent the total number of origin countries ($NB_{OC}$) as a new feature.

Because our model predicts the next year’s ER response in each country solely based on its features, it is agnostic to the name of the countries, and countries are not used as exploratory feature. on the other hand, globally, some regions are more conflict–ridden and a crisis in one country effects other countries in the region. As often the resident country of refugees is in the same geographical neighborhood of the origin country, using residence countries’ geographical regions can perform as a good indicator of the ER response. Therefore, the region each country is located
in can contribute to a better estimation of the ER response. We created a factor that represents the geographical region of the residence country (“RES_(name of the country”).

### 3.3.4 Exploratory Analysis of ER Response and Explanatory Variables

Emergency relief response is both uncertain and volatile throughout the years across different countries. Figure 3-2 shows the amount of dollars spent for emergency relief response by UNHCR from 2011 to 2016. Countries have very different trends in their ER response throughout the years. For example, Afghanistan had ER response only in 2015. Many countries became the residence country for refugees after unset of emergency in others. Examples are Greece in 2015 after the Syrian refugee crisis and South Sudan due to refugees fleeing Sudan. Even with countries experiencing conflicts or receiving ER response every year the amount of ER response fluctuates drastically from one year to another. Examples are Ethiopia, Democratic Republic of the Congo, Iraq, Jordan, Rwanda, Tanzania, Turkey, and Uganda. Forecasting ER response requires a systematic approach that can take into account the uncertainty of requiring ER response as well as considering the fluctuations in its amounts.

Emergency response data —our response variable— is sparse by nature as many countries have zero ER response while some other receive large amounts of ER response. Figure 3-3 illustrates the frequency distribution of ER response Expenditures Between 2011 and 2016.

From countries served by UNHCR, 64% have zero ER response. On the right,
Figure 3-2: UNHCR ER Response Expenditure from 2011 to 2016 in refugees’ residence countries

Figure 3-3: Left: Frequency Distribution of UNHCR ER Response Expenditure from 2011 to 2016, Right: Frequency Distribution of Positive UNHCR ER Response Expenditure from 2011 to 2016

Figure 3-3 shows that even after eliminating the countries with zero ER response from the data, the distribution of ER response is right skewed and the long tail demonstrates that some countries do experience extremely large ER responses. Having a zero-inflated and also right-skewed ER data shows that many countries experience either of the extreme cases: zero or very high response. We observed that in our analysis single-attempt prediction models, trying to simultaneously capture both ex-
tremes, may yield large prediction biases and errors. In section 3.4 we devised a two-step approach suited for predicting ER response from this data.

In section 3.3.3, we created two set of features representing the geographical regions of residence countries and origin countries. To illustrate the importance of including residence regions as one of the features for predicting ER response, we show the amount of ER spent at each region between the years 2011 and 2016 in Figure 3-4. This figure confirms that some regions do experience greater ER response compared to others. For instance, eastern Africa, south-east Asia, and western Asia have the most ER response among all regions. On the other hand, To illustrate the importance of including origin regions, we show refugees movement between different geographical regions in Figure 3-5. Larger circles on the secondary diagonal shows that, most refugees have moved to their neighboring countries within their own geographical region. In addition, larger circles that are not on the secondary diagonal shows that, most refugees have moved to their neighboring regions. For example, most refugees in south-east Asia have moved to southern Asia and eastern Asia while most refugees in northern and middle Africa have moved to eastern Africa.

Exploring the impact of adopted features on ER response and investigating the correlation among features themselves, we created Figures 3-6 and 3-7 to show the pairwise associations between all variables and cluster variables that have relatively stronger associations. In the correlation matrix illustrated in figure 3-6, we used the Pearson, Point Biserial, and Phi correlations to determine the correlation between (1) two continuous variables, (2) a continuous and a binary, and (3) two binary variables, respectively. Continuous variable are depicted in black and binary variables are shown in grey. Larger and darker circles represent greater positive or negative (blue or red) linear relationships for continuous variables, and more frequent co-occurrence
Figure 3-4: UNHCR ER Response Expenditure in refugees’ residence regions

for binary variables. The patterns are intuitive and as expected. For instance, variables such as INFORM risk, vulnerability, and lack of coping capacity demonstrate a less robust country and are negatively correlated with positive economical indicators such as GNI and GPD. To better demonstrate the correlations graphically, we ran a cluster analysis (using hierarchical clustering) to place correlated features closer to each other. Features placed next to each other having larger and darker circles on the diagonal are members of same cluster. The figure reveals five main clusters in our data with the highest internal correlation.

We also look at continuous variables’ rank-based correlations using the Spearman correlation in Figure 3-7 because the Pearson correlation may be low, even when there is a strong association between variables [42].

Figure 3-6 shows that ER response, denoted by ER on the figure axis, does not have large linear relationships with any of the continuous features, except for ER opera-
tion itself with a one–year lag (ERO). However, as can be seen in Figure 3-7, it has a strong association with many of the continuous features such as GNI, OO, refugees assisted, UNHCR refugees, IDPs, BRD, and INFORM risk index.

We also see that source ER and Source OO are negatively associated with origin countries which are more vulnerable to conflicts and have greater assistance rates.
Figure 3-6: Pairwise Association Among Features and ER Response: Pearson correlation for two continuous variable, and Point Biserial correlation for binary and continuous variables, Phi-coefficient for two binary variables

In other words, source _ER and Source _OO deduce information on which countries have the potential to attract refugees and represent countries that are more stable for refugees immigration.
Figure 3-7: Pairwise Association Among Features and ER Response: Spearman ranked-based correlation between two continuous variables

Focusing on exploratory analysis on the data, an important consideration is the possible presence of correlation among the features. This can cause the prediction models to overfit or underfit the data, increase the variance of predictions, and complicate the identification of ER response predictors. In section 3.4.1.1 we introduce an algorithmic step-by-step procedure to alleviate the correlation-caused problems in the data set to an acceptable degree.
3.3.5 Limitations

We end this section with two main caveats. First, ideally, the goal is to predict ER response. However, measuring the actual emergency demand in practice is extremely difficult. We can only use the data on the amounts of funds spent to assist refugees which, in reality, represents the response rather than actual demand. For example, if the amount of dollars spent for emergency relief response in a country is low, it may not necessarily mean low ER response. One possibility is that the un-availability of information for that country causes under-representation in the media and other outlets, and the affected country may not attract donations and funds. As a result, any analysis is vulnerable to incompleteness of information provided. Even thought we adopt a much wider spectrum of features than the existing literature, each country may have unique un-recorded policies and conditions that can cause additional complexities. The data required for addressing this issue at a scale comparable to this study are still not widely available to researchers.(maybe Marianne/Joakim can provide a better and more comprehensive explanation here)

The second caveat is that, we do not answer the separate but important question of how much of each relief item countries will need. Although this paper aims to estimate response levels and identify emergencies, we do not have the right-level detailed data set on the relief items required per country throughout the years. Further, decision makers care about risks of future emergencies and required funds to be allocated to emergency locations particularly in the planning phase. They invest in decision support tools that demonstrate such information and help adapt the current structure in the pre-planning of humanitarian logistics decisions, rather than focusing on the amount of the individual relief item.
3.4 Modeling with an Empirical Strategy: Forecast the Amount of ER Response in Each Country

Having a limited, zero-inflated, and skewed response variable – ER response – disables many state-of-the-art statistical and machine learning models from providing precise predictions in each country, namely because most such methods require a large sample size or focus on minimizing the overall prediction errors. Our goal in this section is to propose a framework for forecasting ER response such that the prediction accuracy at each country is improved along with the overall errors compared to state-of-the-art learning models.

Figure 3-8 illustrates the step by step process of the proposed framework for forecasting ER response in each country. Let the data set that includes all the features collected from UNHCR, WB, UCDP, and INFORM be the “initial feature set”. In the first step, we selected the most relevant features for predicting ER response from the initial feature set through a comprehensive correlation analysis. Before building our model, we develop benchmark models that utilize a state-of-the-art machine learning algorithm. The purpose of developing benchmark models is to track and evaluate the improvements attained by each step of the framework including the preprocessing steps such as feature selection, and also model development techniques, such as model architecture. To this end, one of the most powerful learning algorithms is used to quantify the significance of adopting the framework presented in this work. The so-called eXtreme Gradient Boosting (XGBoost) algorithm was introduced by Chen et al. in 2015 and does not make assumptions about the distribution of data or the error terms [21, 20]. XGBoost reduces prediction bias and uses a regularized
formulation to avoid over-fitting. It also uses subsets of observations and features to simultaneously reduce variance as well. Feeding the selected features from the initial feature set to the XGBoost algorithm leads to the first benchmark model (Benchmark Model A).

Next, we added the set of engineered features, explained in Section 3.3.3, to the initial feature set and selected the most relevant features using the correlation analysis. The selected feature set is then used to create the second benchmark model (Benchmark Model B) using the XGBoost algorithm. We compare the performance of the two benchmark models and examine the added value of including engineered features. Recall that all models in this study use only the training data set for learning and validating the relationship between the selected features and ER response. Consequently, the performance of each model is evaluated by test errors such as MAE, RMSE, and $R^2$ for the regression models. We also plot the predicted versus actual ER response to illustrate the accuracy of individual predictions.

In Figure 3-8, the training module shows that the predictive model for forecasting the amount of ER response is trained in two learning stages. In section 3.2, we explained that predictive modeling literature for zero-inflated variables, such as the ER response, is divided to two main groups which contradict on the assumption that whether zero and positive values are drawn from the same or separate distributions. On the line of thought with same distributions, ER response data would be most suited to Tweedie regression [85] which assumes that the mean response has a Tweedie distribution in a Generalized Linear Model (GLM). Alternatively, Duan et al. [26] proposed a two-part method that is based on the assumption that zero and positive values are drawn from separate distributions. This method uses a binomial
model to predict whether the response variable will adopt zero or positive values ("occurrence") then uses a GLM regression methodology with a log-normal link function to model the observed positive values ("intensity").

As we observed a better predictive accuracy from the two-step model or the occurrence/intensity model, our proposed framework utilizes a two-part logic. In the first part of the training module, the occurrence of non-zero ER response is predicted through training a classification model. This model identifies the countries that will experience a non-zero ER response and is called "Occurrence Model (C1)". In the second part of the training module, the amount of ER response is predicted by a regression model trained on the portion of the training data that has a non-zero ER response. This model is referred to as "Intensity Model (C2)".

In the prediction module shown in the lower-right of the Figure 3-8, we run the occurrence model (C1) on the test data and predict the set of countries that will experience a non-zero ER response in 2016. Next, we run intensity model (C2) on the identified countries in C1 model to predict the amount of ER response. Note that, this means the predicted ER response for the remaining countries is equal to zero. In the following subsections, the building blocks of the Figure 3-8 are discussed in more details.
3.4.1 Feature Selection Based on Correlation Analysis

3.4.1.1 Correlation Analysis of the Initial Feature Set

Strong associations and multicollinearity among features inflate the standard errors of estimated features’ coefficients in a prediction model, degrade the model performance, and complicate the interpretation of estimated coefficients. To include only the relevant features in the prediction models, we used a correlation analysis to identify the relevant features that improve the model performance.

In this study, we aim to provide humanitarian organizations with the best set of features and information arrangement format for input to the ER response forecasting models. In this context, we intend to identify the best way of incorporating refugees’ origin information. In Table 3.3, we showed refugees’ origin information by two groups of binary variables (ORG_(name of the country) and ORG_(name of the
Theses two groups represent refugees’ origin country and the geographical region of their origin country, respectively. In section 3.4.3.2, we will deep dive into the importance of including refugees’ origin information. We will also compare the performance of models when including either origin countries or the origin regions. Therefore, the two binary groups of refugees’ origin information, ORG_ (name of the country) and ORG_ (name of the region), are excluded from the feature selection module in framework depicted in Figure 3-8.

Figure 3-6 shows the pairwise associations between all features except for the refugees’ origin information. Additionally, the associations between continuous features are investigated through the Spearman rank-based correlation illustrated in Figure 3-7. Besides correlations, we calculate the Variance Inflation Factor (VIF) associated with each feature.

For the feature with the highest VIF, we identified the group of features that are highly correlated with it (0.75), using the correlation matrix in Figure 3-6. In every highly correlated group, we eliminated the feature that has the least association with the ER response based both correlation matrices. If two or more features are observed to have equal or similar associations (0.05 difference) with the ER response, the feature with the least correlation with all members of the group is eliminated. After each elimination, the VIF is re-calculated and the same process is repeated until the VIF of all features is less than the conventionally accepted value of 5.

It might be counter-intuitive to eliminate the feature with the least correlation with other features in the highly-correlated group, while selecting features that reduce multicollinearity. Note that, feature selection also aims to maintain features that are
representative of eliminated ones, particularly those that are correlated with response variable. Let us explain with an example. INFORM RISK has the largest VIF value among all features. From correlation matrix in Figure 3-6, we find the group of features that are highly correlated with INFORM RISK include Vulnerability, Lack of Coping Capacity, and Hazard and Exposure. The Pearson correlation between the ER response and all members of the group is negligible, however, the Spearman matrix shows that the least association with ER response belongs to Hazard and Exposure. Therefore, Hazard and Exposure is eliminated from the features. In the next iteration, INFORMS Risk remains as the feature with the highest VIF. We eliminated Lack of Coping Capacity as it has the least association with ER response. In the third iteration, INFORM Risk remains as the variable with the largest VIF. The Spearman matrix shows that ER response is almost equivalently associated with INFORM RISK and Vulnerability. Therefore, we focus on the associations among all four features of the highly correlated group. INFORMS RISK has the largest pairwise correlation with each feature in this group separately. As a result, Vulnerability is eliminated and INFORM RISK is kept as the representative feature of this group. Recall that, INFORM Risk is actually a combination of the other three features in this group and having this process selecting INFORM Risk shows that we correctly identified INFORM risk as feature that explains variations in this group.

By the end of this process, we filtered out Hazard and Exposure, Lack of Coping Capacity, Vulnerability, Refugees Assisted, and Labor Force. This process reduced the VIF from 60 to 2.5. To further reduce the effects of correlations, model-specific techniques will be used.
3.4.1.2 Correlation Analysis with Additional Factors

We now add the engineered problem-specific features to the initial feature set. We repeat our feature selection methodology using correlation analysis explained in section 3.4.1.1 with the inclusion of the engineered factors. Similar to the correlation analysis of the initial feature set, origin countries and origin regions are excluded from the elimination process. After calculating the VIF associated with each of the features, all the added engineered factors showed a VIF < 5. As a result, the same variables as in section 3.4.1.1 are eliminated. This step reduced the VIF from 60 to 4.

So far, we created two data sets with VIF < 5. In section 3.4.2, we will use them to create the two state-of-the-art machine learning models (Benchmark Model A and B). These models are benchmarks for evaluating the performance of our proposed framework.

3.4.2 Benchmark Models and Their Comparison

Using the selected features from the initial feature set and the origin countries, we developed the benchmark model A using the XGBoost algorithm to predict the ER response. We used a 5-fold cross-validation with 10 repeated samplings and used a grid search method to optimize the tuning parameters. We selected the model that yields the least out of the bag error. Our performance metric for choosing the best model is the RMSE as RMSE penalizes large errors more than MAE. This is in line with our purpose of capturing extreme cases where ER response is zero or has large positive values. Figure 3-9 shows the ER response values predicted by benchmark model A for the test data set.
Next, we created a new benchmark model (B) with the addition of some of the engineered factors. Among the engineered factors is the origin regions and the results of inclusion of either origin countries or origin regions are compared. Moreover, “Number of Origin Countries” (NB_OC) is only included when using the origin region. This model uses the same algorithm, grid search, and cross-validation steps as in benchmark model A. We find that the best model includes selected features from the correlation analysis in section 3.4.1.2 and origin countries rather than the origin regions.

Table 3.4 shows the performance metrics of benchmark models A and B on the test data. Benchmark model B, with a slightly better performance, makes an average error of $941,695 in its predictions and leads to $R^2$ equal to 25%.

Figure 3-11 compares the predicted values from the two benchmark models A and B with the actual ER response. The y-axis shows the predicted ER and the x-axis
Figure 3-10: Predictions from the Benchmark Model B: Actual versus predicted ER response for all countries

Table 3.4: Performance Metrics: Comparison of benchmark models (A) and (B)

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Benchmark Model A</th>
<th>Benchmark Model B</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td>22%</td>
<td>25%</td>
</tr>
<tr>
<td>RMSE</td>
<td>1,934,650</td>
<td>1,893,376</td>
</tr>
<tr>
<td>MAE</td>
<td>985,393</td>
<td>941,695</td>
</tr>
</tbody>
</table>

shows the actual ER response. The dashed red diagonal line represent the perfect alignment of predicted values with the actual ER response. In both figures, the points with the largest errors happen at the two extremes of zero or very high actual ER values. This shows the shortcomings of standard machine learning methods for predicting extreme cases especially when the data is zero-inflated and skewed with long tails.
3.4.3 Training Module: Two-part Learning Algorithm for Forecasting ER Response

Our proposed model follows the two-part model logic of Duan et al. (1983). However, unlike Duan et al., we do not restrict the classification decisions to a GLM model with a binomial distribution.

The first part of the proposed model refers to a classification model (occurrence model C1) that identifies the class of each country, denoted by $\phi_i \in \{0, 1\}$. If a positive ER response is incurred in country $i$, then $\phi_i = 1$ and is said to have a positive class. Inversely, if a zero ER response is incurred, then $\phi_i = 0$ and it has a negative class. Next, a GLM regression model (intensity model C2) predicts the amount of ER response in countries where a positive ER value is expected to occur.

Equation 3.3 shows our developed two-part model where two different binary classification models, represented in (i) and (iii), are combined to construct the first stage occurrence model C1. Part (iii) of Equation 3.3 represents the second stage
intensity model C2, and models the mean ER response of countries using a Gamma regression. Recall that in section 3.3.2, we showed that our goal is to find a mapping function \( f(x^t) \) that explains the relationship between features of year \( t \) with ER response in year \( t+1 \). Let \( X^t \) be the \((n \times p)\) matrix of features where \( p \) is the number of features selected for each model, and \( \beta_{js} \) be the coefficients of feature \( j \) in model \( s \) where \( j \in \{1, \ldots, p\} \), \( s = \{(i), (ii), (iii)\} \).

\[
\begin{align*}
f(x^t_i) = & \begin{cases} 
\min_{\beta_{(i):i} \epsilon_i} C \left[ \sum_{i=1}^{n} \max[0, 1 - \text{ER}_{i}^{\text{bin},t+1}k(x^t_i)] \right] & (i) \\
\text{s.t } \text{ER}_{i}^{\text{bin},t+1}(k^{(t)}_{(i)}) \geq M(1 - \epsilon_i), \sum_i \epsilon_i \leq v, \epsilon_i \geq 0 \\
\logit(p^{t+1}^i) = \sum_j^p x^t_{ij} \beta_{(ii)} & (ii) \\
\log(\text{ER}_{i}^{t+1}|(\text{ER}_{i}^{t+1} > 0) = \sum_{j=1}^{p} x^t_{ij} \beta_{(iii)} & (iii)
\end{cases}
\end{align*}
\]

Part (i) of Equation 3.3 presents the Support Vector Machines (SVM) classification formulation [75]. In this formulation, \( C \) is a tuning parameter known as cost, \( M \) is a predefined constant large number, and \( v \) is a predefined constant error term. The outcome of the SVM model is the function \( k(x^t_i) \) which determines whether a positive ER response is incurred in country \( i \). This function is typically referred to as the decision boundary function and is translated into classes \( \phi^{SVM} \) using Equation 3.4.

\[
\phi_i = \begin{cases} 
\phi_i = 1 & if k(x^t_i) > 0 \\
\phi_i = 0 & if k(x^t_i) < 0
\end{cases}
\]
Part (ii) of Equation 3.3 represents the Logistic Regression (LR) model. The outcome of LR model is the probability that ER response in country $i$ adopts a positive value ($p_i^t = Pr[ER_i^{t+1} > 0]$). Typically, the LR probabilities $p_i^{t+1}$ are translated into classes by setting a threshold. Classes derived by the LR model are referred as $\phi^{LR}$.

The SVM model constructs the basis of our occurrence predictions since it directly identifies a class for each country as opposed to the LR model, where a choice of threshold is required. In this analysis, the SVM model also provides significantly more accurate results than the LR model, which will be presented in section 3.4.3.2. However, we still aim to use the LR classes and probabilities as a safety valve to amend predictions only when the probability is outside the acceptable interval. Since LR and SVM use different classification techniques, each can perform better for particular cases. As a result, eliminating one in favor of the other indiscriminately can lead to larger number of missclassifications.

The outcomes of the two classification models are combined using Equation 3.5 to calculate adjusted SVM outcomes ($k_{adj}(x_i^t)$). In this equation, $L$ denotes the lower threshold and $U$ is the upper threshold of the acceptable interval. The adjusted values are now fed to Equation 3.4 to calculate the classes of the occurrence model $C1$ as the first part of the two-part learning model.

$$k_{adj}(x_i^t) = \begin{cases} k(x_i^t) & L \leq p_i^{LR} \leq U \\ -k(x_i^t) & \phi_i^{SV} \neq \phi_i^{LR} \text{ and } p_i^{LR} < L \text{ or } p_i^{LR} > U, \end{cases}$$  \hspace{1cm} (3.5)$$

Section 3.4.3.2 illustrates the results of models (i), (ii) and investigates the impact of features on accuracy of classification decisions.
Next, we model the mean ER response in part (iii) of Equation 3.3, called the intensity model C2, given that positive ER responses were actually incurred using a generalized Gamma regression model. The intensity model C2 assumes that the mean of the positive ER responses has a Gamma distribution $\bar{ER}|_{(ER_i>0)} \sim \text{Gamma}(\theta, \nu)$, where $\mu_{ER_i} = \frac{\nu}{\theta}$.

In the proposed two-part model shown in Equation 3.3, we take advantage of the fact that the features used in the two learning stages need not be the same. Also, since at each stage we tune the models to predict either the occurrence or the intensity of positive ER response, the models’ prediction performance is enhanced. In fact, in the second learning stage, we show that a much reduced feature set (compared to the first learning stage) suffices to estimate the intensity of ER response. Due to a reduced number of features in the Gamma regression model, the interpretability of feature coefficients is increased.

Section 3.4.3.1 explains the steps of building occurrence model C1 which classifies countries whose ER response is positive. Section 3.4.3.3 explains details of creating the GLM regression model.

### 3.4.3.1 First Learning Stage: ER Response Occurrence Model C1

Countries with zero ER response are called “Negative Class” countries, and countries with non-zero ER response are called “Positive Class” countries. In building the occurrence model C1, we aim to maximize the number of correctly identified positive class countries. In other words, when predicting the positive class countries, our metric of interest cares about minimizing the number of “misses”. We also want to maximize the number of correctly identified negative class countries. That is, minimizing the “false alarms”. As a results, minimizing total number of misclassifications
(or maximizing the accuracy of the prediction) summarizes our metric choice. Note that, when dealing with imbalanced data, namely when there are disproportionate ratio of countries in each class, developed model may suffer by tuning the parameters using accuracy. However, in our analysis, the ratio of positive to negative classes is 3 to 1. This ratio is not severe enough to negatively impact our model selection decisions. In fact, using accuracy as the metric led to a better balance of misses and false alarms, which is of substantial importance for our purpose, compared to a model tuned using an imbalanced data metric such as the F1-score. False alarm countries undergo a regression model in the next section, where a positive value is predicted as their ER value. Therefore, having misclassifications being solely comprised of false alarms increases our ultimate prediction error. On the other hand, missing prediction of non-zero countries leads to increases our ultimate prediction error.

Now, recall that we want to answer to the question: which combination of data with the problem-specific factors provides better classification? We used the selected features from the correlation analysis in section 3.4.1.2 and added source_ER, source_oo, and Refugee_Assistance as ENG, origin countries as ORG.C, origin regions, the number of origin countries as ORG.R, and finally residence regions as RES_R in different combinations. The fixed portion of data is represented with Latin symbol (I).

We compared the performance metric of LR and SVM classification models with each of the feature combinations and selected the combination that minimizes the total number of misclassifications. The precision, recall, F1-score and the accuracy of each model is recorded for further analysis explained in section 3.4.3.2.
For the LR model, we used a Ridge regularization method to avoid overfitting the model. For model tuning and validation of the LR model, we optimize model parameters—regularization multiplier and threshold—through a grid search and 5-fold cross validation with 10 repeated sampling. The best model has a regularization multiplier of 0.007 and a threshold of 0.44. Classifying the test data using the LR model with a threshold of 0.44, we find a high prediction bias showing a large difference between the average of predictions and the actual average of data classes. To resolve this problem, we used the plat-scaling method to calibrate the probabilities from the LR model.

For the SVM model, we used a Radial Kernel formulation for the $k(x_i^t)$ function, and tune model parameters—Radial Kernel parameter $\gamma$ and cost $C$ in the SVM objective function—through a grid search with 5-fold cross validation and 10 repeated sampling. The best model that minimizes out-of-the-bag misclassifications has a $\gamma$ of 0.048 and a $C$ of 3. After identifying the tuning parameters, we develop a SVM model that uses the entire training data while setting $\gamma$ and $C$ to the determined values. In both methods, the best model and feature sets are selected simultaneously.

Figure 3-12 represents the predicted probability of each country using LR model
the class predicted by SVM model $\phi_i^{t+1}$, and the actual class of ER response $ER_{i,bin,t+1}$. Finally, we revised the predicted SVM classes using the probabilities that the LR model provides based on Equation 3.5 by setting $L$ to 15% and $U$ to 85%. That is, we switched SVM predictions if LR probabilities is in significant disagreement with the predicted class. Figure 3-13 shows the performance of the occurrence model C1 on test data. we switched SVM predictions if LR probabilities is in significant disagreement with the predicted class.

![Figure 3-13: Model Performance of Occurrence Model C1: Number of correct identifications versus missclassification](image)

3.4.3.2 First Learning Stage: Connection to Data

Figure 3-14 compares the LR and SVM classifications when different subsets of the problem-specific factors are added to the selected features in section 3.4.1.2. The
The number of misclassifications from the SVM model is lower than the LR in all cases of different feature combinations. The minimum misclassifications is found where SVM model includes the features selected from the correlation analysis (I) in combination with origin countries and resident regions (I_ORG.C_RES.R). It is interesting to see that when origin countries are not used in the SVM model, its impact is almost entirely compensated by including the origin regions and some engineered factors (I_ENG_ORG.R_RES.R). This combination is, in fact, the feature set that minimizes the misclassification error in the LR model. This feature set includes the selected features (I) and source OO, source ER, number of origin countries, refugee assistance rate, origin regions, and resident regions.

Since the SVM model led to less misclassifications, we use this model to further investigate the impact of origin countries, origin regions, and resident regions on the number of misses and false alarms. In Figure 3-15, we compare the impact of including origin countries, origin regions, and resident regions on the number of

Figure 3-14: Model Performance of LR versus SVM Model: Total number of misclassification under different feature combinations
misses and false alarms, added individually to the initial set and through different combinations with the engineered factors. We find that including resident country’s geographical region reduces missed positive class countries compared to including information about the origins. On the other hand, we find that including refugees origin information with the country or region, results in less False alarms compared to resident country’s geographical region.

![Number of Missed Countries and False Alarms](image)

Figure 3-15: Model Performance of SVM Model: Number of missed countries and false alarms under different feature combinations

### 3.4.3.3 Second Learning Stage: ER Response Intensity Model C2

Using countries with non-zero ER response in the data, we train a GLM model to estimate the value of ER response assuming a Gamma distribution, where the probability of predicting ER values equal to zero is zero. We link function used in C2 model is Logarithmic.

Since using regularization can bias the point estimate of coefficients while reducing their variance, we develop a technique to select relevant features without directly
using regularization. To this end, first, we develop a regularized Gamma regression model (Lasso) using the fixed portion of data (I), Source_OO, Source_ER, number of origin countries, Refugee_Assistance, origin regions, and resident regions. Recall that this is the feature set that yielded the best result for LR model, highlighted in orange in Figure 3-14, which is also a GLM model that uses a Sigmoid link function. For each shrunken feature in the regularized model, we identify the group of features that are most correlated with each other using Pearson correlation. In this group, we eliminate variables base on the process explained in the feature selection using correlation analysis explained in section 3.4.1.1. We end up eliminating OO, IDPs, GNI, GDP_Growth, Unemployment, Refugee Assistance, Source OO, Origin regions. This step reduces the VIF further down to 1.6.

3.4.3.4 Second Learning Stage: Connection to Data

Validating the developed Gamma regression model on the non-zero ER response portion of the test data leads to $R^2$ equal to 65%, MAE equal to $958,025$, and RMSE equal to $1,199,584$. Figure 3-16 compares actual ER values and predicted values for countries with non-zero ER response in 2016. It is clear that proposed model captures countries with significantly large ER response and countries with smaller ER requirement. Recall 3-10 where benchmark model B predicted ER values no more than 4 million, and did not identify countries with significantly high or low ER requirement. One of the questions we want to answer with our analysis is: what are the main predictors of ER response? The answer to this question is in two fold. Figure 3-17 indicates the relative importance of each factor, with a collective effect perspective, from the SVM classification model. Having multicollinearity reduced in section 3.4.3.3 without adoption of a regularization method, the direction and
relative importance of estimated feature coefficients provide insightful and reliable information. Economic factors such as GNI, GDP growth, GDP and refugees origin countries and countries’ population are the top 5 factors positively associated with occurrence of ER response. On the other hand, dollars spent on ongoing and emergency operations, number of IDPs, Returned IDPs, and the INFORMS Risk factor are the 5 top predictors negatively associated with having positive ER response in a country. Figure 3-18 indicates the relative importance of each factor, with a collective effect perspective, from the Gamma regression model.

Given that a country will experience ER response, country’s population, number of people categorized as others of Concern, INFORM Risk, Residence Region - countries in south east Asia, south Europe, western Asia - positively impact the amount of ER response. Residence Region - countries in northern America, northern Europe, western Europe -, number of returned IDPs, UNHCR refugees, and number of BRDs negatively impact the amount of ER response.
In terms of UNHCR’s population of concern, in both classification and regression models others of concern is positively associated with ER response and returned IDPs, UNHCR refugees, BRD are negatively associated with ER response.

Figure 3-17: Effect of Features on Occurrence of a positive ER Response from the SVM Model
3.4.4 Prediction Module

3.4.4.1 Predict Occurrence of Positive ER Responses using C1 Occurrence Model (C3)

Feeding Test data to the occurrence model C1, we classified a total of 40 countries as positive class countries. ER response of the remaining 54 countries in 2016 is set to zero. With accuracy of 84%, the classification model captured 86% (recall) of countries with have ER response. 75% of predicted Positives were actually positive (precision), and 83% of predicted Negatives were actually negative (specificity).
From figure 3-12, we see that both LR and SVM models classified Pakistan, Turkey, Lebanon, Cote d’Ivoire, Afghanistan, Myanmar, Mali, Eritrea, and Congo as positive class countries, while their actual ER response in 2016 was zero. Figure 3.5 shows the actual ER and OO value of false alarms across 2011 to 2016. These countries except Hungary and Pakistan, either have experienced large ER values in the previous years (Afghanistan), or have been constantly experiencing emergency situations throughout the years (Cote d’Ivoire, Lebanon, Mali, Myanmar, and Turkey). Table 3.6 shows the actual ER and OO value of missed countries. All of these countries have either small amounts of ER response in 2016 (Albania, Bangladesh, Slovenia, and Zambia), or had ER response for the first time in 2016 (Ecuador).

Table 3.5: Actual versus Predicted ER Response and the Amounts of Past ER and OO Expenditures in False Alarm Countries

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Afghanistan</td>
<td>0.8</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>7,080,830</td>
<td>301,454</td>
<td>2,095,452</td>
<td>7,421,612</td>
<td>6,391,041</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3,588,127</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cote d’Ivoire</td>
<td>0.83</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>657,109</td>
<td>0</td>
<td>56,496</td>
<td>0</td>
<td>41,510</td>
<td>0</td>
<td>182,038</td>
<td>185,530</td>
<td>95,208</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eritrea</td>
<td>0.76</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>105,213</td>
<td>0</td>
<td>0</td>
<td>1,206,625</td>
<td>297,075</td>
<td>168,882</td>
<td>205,104</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungary</td>
<td>0.4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>985</td>
<td>0</td>
<td>0</td>
<td>390,685</td>
<td>75,451</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lebanon</td>
<td>0.84</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>6,500</td>
<td>1,202,700</td>
<td>3,191,934</td>
<td>3,390,066</td>
<td>1,181,075</td>
<td>3,608,883</td>
<td>2,205,788</td>
<td>1,249,130</td>
<td>3,390,685</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mali</td>
<td>0.83</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>56,281</td>
<td>8,364</td>
<td>25,639</td>
<td>0</td>
<td>174,393</td>
<td>439,093</td>
<td>1,436,068</td>
<td>9,1903</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myanmar</td>
<td>0.67</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1,002,431</td>
<td>460,887</td>
<td>20,456</td>
<td>430,795</td>
<td>0</td>
<td>5,004,420</td>
<td>2,470,751</td>
<td>1,648,772</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pakistan</td>
<td>0.90</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>14,034,727</td>
<td>6,732,620</td>
<td>760,948</td>
<td>7,066,321</td>
<td>4,463,648</td>
<td>4,525,350</td>
<td>4,525,350</td>
<td>4,463,648</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turkey</td>
<td>0.84</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>7,849,000</td>
<td>1,026,553</td>
<td>13,565,576</td>
<td>7,555,500</td>
<td>1,252,807</td>
<td>7,721,006</td>
<td>2,317,237</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.6: Actual versus Predicted ER Response and the Amounts of Past ER and OO Expenditures in Missed Countries

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Albania</td>
<td>0.19</td>
<td>0</td>
<td>1</td>
<td>80,230</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2,098</td>
<td>0</td>
<td>26,065</td>
<td>18,230</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bangladesh</td>
<td>0.27</td>
<td>0</td>
<td>1</td>
<td>179,736</td>
<td>0</td>
<td>127,623</td>
<td>29,960</td>
<td>37,344</td>
<td>0</td>
<td>794,548</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>179,736</td>
<td></td>
</tr>
<tr>
<td>Ecuador</td>
<td>0.24</td>
<td>0</td>
<td>1</td>
<td>1,110,810</td>
<td>0</td>
<td>1,110,810</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1,110,810</td>
<td></td>
</tr>
<tr>
<td>Slovenia</td>
<td>0.26</td>
<td>0</td>
<td>1</td>
<td>80,280</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>282,555</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>282,555</td>
<td></td>
</tr>
<tr>
<td>Zambia</td>
<td>0.31</td>
<td>0</td>
<td>1</td>
<td>88,832</td>
<td>0</td>
<td>120,156</td>
<td>29,666</td>
<td>44,993</td>
<td>2,315</td>
<td>60,954</td>
<td>6,700</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>88,832</td>
<td></td>
</tr>
</tbody>
</table>

128
3.4.4.2 Predict the Amount of ER response Using C2 Regression Model (C4)

We predict required ER response for the 40 positive class countries identified in the previous section. Setting ER response of the negative class countries to zero, Figure 3-19 shows the final predictions of the proposed model. We predicted 2016 ER values of each country with average error of $389,694. Recall that the best benchmark model mean absolute error of $941,695. This is 58% reduction in mean errors compared to benchmark model B. $R^2$ of the proposed model is 70% on test data, whereas state-of-the-art benchmark model B had $R^2$ of 25% in this unseen portion of data.

Figure 3-19: Predictions from the Proposed Two-Part Learning Algorithm: Actual versus predicted ER response for all countries
Table 3.7: Performance Metrics: comparison of proposed model versus benchmark models

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Benchmark Model A</th>
<th>Benchmark Model B</th>
<th>Benchmark Model C</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td>22%</td>
<td>25%</td>
<td>70%</td>
</tr>
<tr>
<td>RMSE</td>
<td>1,934,650</td>
<td>1,893,376</td>
<td>1,199,584</td>
</tr>
<tr>
<td>MAE</td>
<td>985,393</td>
<td>941,695</td>
<td>389,694</td>
</tr>
</tbody>
</table>

3.5 Additional Analysis: Forecast the Severity Levels of ER Response in Each Country

Our goal in this section is to provide probabilistic forecasts indicating the likelihood of experiencing low, medium, and high ER responses in each country and to use them to improve and analyze results from the two-part learning model. From a practical point of view, any ER amounts below $100,000 is considered negligible and above 5M is treated as extremely large. We divided the ER response range into 3 intervals representing low, medium, and high classes of ER severity. Table 3.8 shows the amount of ER response associated with each class. This classification is drawn from our collaboration with UNHCR. In this classification, zero or negligible ER responses ($< 100,000$) are grouped in the “Low” ER severity level and extremely large ER responses ($> 5M$) are classified as the “High” ER severity level. Non-extreme positive ER amounts are classified as “Medium”.

Table 3.8: Ranges Associated with Classes of ER Severity

<table>
<thead>
<tr>
<th>Class</th>
<th>Range [USD]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>&lt;100,000</td>
</tr>
<tr>
<td>Medium</td>
<td>100,000—5,000,000</td>
</tr>
<tr>
<td>High</td>
<td>&gt;5,000,000</td>
</tr>
</tbody>
</table>

To predict the probability distribution of the ER response, a multi-class logistic regression model (MC LR model) is trained on the features along with the proba-
bilities and classes generated in the occurrence model developed in section 3.4.3.1. In the MC LR model, the combination of features that minimized the misclassification error of the binary LR model, highlighted in Orange in Figure 3-14, is utilized. Additionally, we added the fitted probabilities generated from the binary LR model and the fitted classes produced by the SVM model to the feature set of the MC LR model. With this addition, we improve the performance of the MC LR model as it learns cases not correctly captured by the occurrence model as well as those correctly identified.

Figure 3-21 shows the predicted probability distribution of ER from the MC LR model. If the classes with the largest probability is selected as the predicted ER severity level, then Figure 3-20 shows the number of correctly identified observations per class. From 65, 25, and 4 countries with actual ER severity levels of low, medium, and high, the MC LR model identified 55, 21, and 0 as such, demonstrating an accuracy of 85%, 84%, and 0%, respectively. The accuracy of the model for the three ER severity classes show that the MC LR model is mostly reliable for detecting low and medium severity situations.

In section 3.4, we proposed a framework for an exact value estimation of ER response per country and proposed a two-part learning algorithm that yielded an unprecedented $R^2$ of 70% and a mean error of $\$389,694$ on the unseen portion of data. The important question to answer here is what insights or pieces of information can the newly introduced probabilistic forecasting method (MC LR model) tell us that the exact forecasting method (two-part learning algorithm) lack?

Let us point to Table 3.9. This table compares the actual ER response severity classes with the predicted severity classes from the two-part learning model and MC LR model. This table shows only the countries whose predicted class in MC
LR model and the proposed two-part learning model are not aligned. In all other cases, the two models depict similar pictures. As noted above, the MC LR model is most reliable when the predicted severity class is low or medium. In such cases, the discrepancy with the two-part learning model can guide us if the our expected ER value is over/under-estimated. However, driving conclusions based on the class with the largest probability does not use the capabilities of this model to the fullest. One needs to pay attention to the distribution probabilities across different classes.

In Figure 3-21, some countries’ probability distribution is spread out across two or three classes, while others retain mostly one-class distributions. For instance, the MC LR model provides a one-class distribution over the low and medium class for the countries of Eritrea, Pakistan, Somalia, and Yemen. This model tells us that the two-part learning model has over-estimated the ER response value for these countries. Therefore, we can modify our ER value estimations from the two-part learning model for a closer value to the lower end of the predicted class range for these countries.

On the other hand, if the probabilities of two or more classes are similar, then a more conservative approach is advised as we cannot recommend modifications to the expected ER value with the same level of certainty as a concentrated probability distribution would have suggested. The decision maker shall make the judgement based on the available information. For example, the probability distribution of South Africa is dispersed across the three classes. The MC LR model has a 34% probability of class low and 49% of extremely high ER. As this distribution is spread across the three classes and also Since the MC LR model is more reliable for detecting the low and medium class, we do not recommend any modifications to the predicted zero ER response from the two-part learning model for South Africa. Turkey with 51% probability for low ER severity level and total of 49% probability for medium
and high ER severity levels has a dispersed probability distribution. Although the two-part learning model predicts a medium ER severity level for Turkey, one can modify the predicted ER response to the lower ends of medium ER severity range or cautiously refrain from making any modifications.

Table 3.9: Comparison of Actual ER Severity Classes with Predicted Severity Classes from the Two-Part Learning Model and MC LR Model in Countries with miss-aligned predictions.

<table>
<thead>
<tr>
<th>Country</th>
<th>Actual 2016 ER</th>
<th>Predicted ER from Two-Part Learning Model</th>
<th>Predicted ER from MC LR Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Value Class</td>
<td>Value Class</td>
<td>Low</td>
</tr>
<tr>
<td>Eritrea</td>
<td>0 Low</td>
<td>583,891 Mid</td>
<td>0.57</td>
</tr>
<tr>
<td>Pakistan</td>
<td>0 Low</td>
<td>119,451 Mid</td>
<td>0.79</td>
</tr>
<tr>
<td>Somalia</td>
<td>871,206 Mid</td>
<td>1,529,972 Mid</td>
<td>0.86</td>
</tr>
<tr>
<td>South Africa</td>
<td>0 Low</td>
<td>0 Low</td>
<td>0.34</td>
</tr>
<tr>
<td>Syria</td>
<td>232,570 Mid</td>
<td>2,064,821 Mid</td>
<td>0.04</td>
</tr>
<tr>
<td>Turkey</td>
<td>0 Low</td>
<td>1,473,863 Mid</td>
<td>0.51</td>
</tr>
<tr>
<td>Yemen</td>
<td>6,874,196 High</td>
<td>7,093,902 High</td>
<td>0.06</td>
</tr>
</tbody>
</table>

In summary, forecasting probability distribution of ER severity levels as a complementary step adds additional information to the exact forecasting methods. This is particularly useful in studies with limited response data. It is important to note that decision making, under-writing, and policies and adopted guidelines in humanitarian sector change by country and situation. Having this in mind, such an additional piece of information for decision makers can facilitate adopting changes in each particular.
3.6 Insights, Recommendations, and Additional Exploration of Results

Our analysis reveals that the proposed framework, unlike the existing statistical learning and state-of-the-art machine learning models, captures the both extreme ER situations where countries experience very large positive or zero emergency relief responses. Recalling the categorization of the existing research methods in section 3.2, we found that using a structure were zeros and positive values are drawn from different distributions is a more appropriate approach for zero-inflated skewed data of refugees’ emergency response. Using the two-part logic, the proposed and developed model significantly improved our prediction power from $R^2$ of 25% to 70%, increased model interpretability by reducing the number of features in models, and identified the necessary features for forecasting the ER response.
Moreover, we found that forecasting the probability distribution of countries’ ER severity levels provide valuable information that an exact forecasting method might miss. Our analysis also reveals that refugees resident country’s geographical region and origin countries or region are key predictors for identifying countries that will experience ER response. Further, we found that including resident country’s geographical region helps reduce the number of missed countries. In addition, including information about the origin locations reduced the number of false alarms. The takeaway is that if humanitarian organizations aim to minimize the missclassification errors and identify countries with positive future ER response, they should incorporate refugees’ residence and origin information in their forecasting models.

We also identified that economic factors such as GNI, GPD growth, GDP, and population are positively associated with the occurrence of ER response. This suggests that refugees move to larger, more populated countries with better economic situations. Further, dollars spent on ongoing and emergency operations are negatively associated with a country experiencing ER response. Thus, countries that have been the destination of refugees for a while experience less refugee influx in the upcoming year. This is in line with the common practices of HOs, which underwrites future demands as ongoing operations in the next year, if a country is already has existing ER operations.(Mariane/Joakim may explain)

The negative association between the the number of IDPs, returned IDPs, and BRD with experiencing a positive ER response suggests larger number of deaths and internal displaced people results in less refugees that would leave their country and require ER assistance.

Amongst UNHCR’s population of concern, the number of asylum seekers and others of concern in a country are positively associated with occurrence of ER response. On the contrary, the number of stateless persons and UNHCR refugees are
negatively associated with having a positive ER response in a country. This is in line with the domain knowledge. (Mariane/Joakim may explain)

In terms of the amount and of ER response, our analysis reveals that, from the countries who experience ER response, countries in southeast Asia, southern Europe, and western Asia are primarily associated with greater ER values.

Furthermore, country’s population, number of people categorized as others of concern by UNHCR, and the INFORM risk index are key predictors of larger ER response values. However, the number of returned IDPs, UNHCR Refugees, and BRDs in countries with positive ER response negatively impact the amount of ER response. That is, factors indicating additional number of refugees in countries with positive ER increase the amount of ER response and factors indicating less number of refugees decrease the amount of ER response. (Mariane/Joakim may come up with additional specific reasoning/example)

Recall that we reduced the feature set used in the regression model. Having many economical, geographical, institutional, and employment factors being eliminated from the regression model in section 3.4.3.3, suggest that INFORM risk index explains a large spectrum of information. However, INFORM risk index only suffices when countries with positive ER response are known.

We also confirm that countries in northern America, northern Europe, and western Europe are primarily associated with less amounts of ER response.
3.7 Conclusion

In this work we compared the performance of existing statistical learning and state-of-the-art machine learning models to predict zero-inflated skewed emergency demand. We show that using modeling techniques with the right approach and processing steps can outperform state-of-the-art methods that systematically use ensemble and adaptive learning even with hard-to-handle data challenges and limitations. The framework developed here is easy to understand and interpret, and yet outperforms state-of-the-art machine learning models, provides unprecedentedly accurate predictions, and helps identify the main predictors of ER response.

Our analysis focuses on learning the relationship between countries’ characteristics with positive ER occurrences, and explaining the impact of those characteristics in ER amounts. In this process, to avoid the problems caused by correlation amongst features, we developed a practical step-by-step algorithmic procedure that reduces multi-collinearity to an acceptable level as explained in section 3.4.1.1. Additionally, to avoid disadvantages of using regularization methods such as biased coefficient estimates and arbitrary feature choices, we introduced an alternate technique explained in section 3.4.3.3.

This study uses multiple data sources, each contributing to data in various aspects. We incorporate a large collection of economical, societal, political, geographical, environmental, and infrastructural factors, as well as employment, human rights, refugee movement, and many other pieces of information related to countries across different years. As a result, we believe our results is practical and have broader impact than the published studies in this area.
Under-writing, adjusting strategies, policies, and adopted guidelines in humanitarian sector can follow different procedure or can change by country, crisis type, and situation. The ultimate goal is to help decision makers with as much information as it is evident in the data to understand refugee behavior, predict possible outcomes, estimate losses of different strategies, identify challenges in the process, and estimate the impact of decisions made. Knowing that decision maker are informed about day-to-day operations and crisis in regions, they can easily understand and link generated information from our analysis and exploit them in planning and management operations.

The emergency response operations for refugees need more emphasis and deeper analysis in different parts of the supply chain, and in particular, the forecasting part. We extensively explained the limitations of our study and areas lacking enough input information (data) (Section 3.3.5), for which there is a great potential to add important insights to the practice of humanitarian emergency response in future.
<table>
<thead>
<tr>
<th>Countries</th>
<th>Predicted Probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afghanistan</td>
<td>0.1</td>
</tr>
<tr>
<td>Albania</td>
<td>0.89</td>
</tr>
<tr>
<td>Algeria</td>
<td>0.85</td>
</tr>
<tr>
<td>Angola</td>
<td>0.92</td>
</tr>
<tr>
<td>Armenia</td>
<td>0.9</td>
</tr>
<tr>
<td>Australia</td>
<td>0.93</td>
</tr>
<tr>
<td>Azerbaijan</td>
<td>0.9</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>0.93</td>
</tr>
<tr>
<td>Belgium</td>
<td>0.93</td>
</tr>
<tr>
<td>Benin</td>
<td>0.93</td>
</tr>
<tr>
<td>Bosnia and Herzegovina</td>
<td>0.87</td>
</tr>
<tr>
<td>Botswana</td>
<td>0.93</td>
</tr>
<tr>
<td>Burkina Faso</td>
<td>0.93</td>
</tr>
<tr>
<td>Burundi</td>
<td>0.24</td>
</tr>
<tr>
<td>Cambodia</td>
<td>0.93</td>
</tr>
<tr>
<td>Cameroon</td>
<td>0.14</td>
</tr>
<tr>
<td>Canada</td>
<td>0.93</td>
</tr>
<tr>
<td>CAI</td>
<td>0.19</td>
</tr>
<tr>
<td>Chad</td>
<td>0.17</td>
</tr>
<tr>
<td>China</td>
<td>0.08</td>
</tr>
<tr>
<td>Colombia</td>
<td>0.85</td>
</tr>
<tr>
<td>Congo</td>
<td>0.37</td>
</tr>
<tr>
<td>Cote d'Ivoire</td>
<td>0.25</td>
</tr>
<tr>
<td>Croatia</td>
<td>0.36</td>
</tr>
<tr>
<td>Denmark</td>
<td>0.52</td>
</tr>
<tr>
<td>Djibouti</td>
<td>0.24</td>
</tr>
<tr>
<td>Ecuador</td>
<td>0.18</td>
</tr>
<tr>
<td>Egypt</td>
<td>0.09</td>
</tr>
<tr>
<td>Eritrea</td>
<td>0.57</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>0.13</td>
</tr>
<tr>
<td>Georgia</td>
<td>0.52</td>
</tr>
<tr>
<td>Ghana</td>
<td>0.32</td>
</tr>
<tr>
<td>Greece</td>
<td>0.2</td>
</tr>
<tr>
<td>Guinea</td>
<td>0.52</td>
</tr>
<tr>
<td>Guinea-Bissau</td>
<td>0.94</td>
</tr>
<tr>
<td>Hungary</td>
<td>0.27</td>
</tr>
<tr>
<td>India</td>
<td>0.69</td>
</tr>
<tr>
<td>Iran</td>
<td>0.89</td>
</tr>
<tr>
<td>Iraq</td>
<td>0.1</td>
</tr>
<tr>
<td>Israel</td>
<td>0.95</td>
</tr>
<tr>
<td>Italy</td>
<td>0.92</td>
</tr>
<tr>
<td>Japan</td>
<td>0.94</td>
</tr>
<tr>
<td>Jordan</td>
<td>0.11</td>
</tr>
<tr>
<td>Korea</td>
<td>0.09</td>
</tr>
<tr>
<td>Kyrgyzstan</td>
<td>0.91</td>
</tr>
<tr>
<td>Lebanon</td>
<td>0.19</td>
</tr>
<tr>
<td>Liberia</td>
<td>0.25</td>
</tr>
<tr>
<td>Libya</td>
<td>0.21</td>
</tr>
<tr>
<td>Macedonia</td>
<td>0.82</td>
</tr>
<tr>
<td>Malawi</td>
<td>0.4</td>
</tr>
<tr>
<td>Malaysia</td>
<td>0.93</td>
</tr>
<tr>
<td>Mali</td>
<td>0.3</td>
</tr>
<tr>
<td>Mauritania</td>
<td>0.21</td>
</tr>
<tr>
<td>Moldova</td>
<td>0.92</td>
</tr>
<tr>
<td>Montenegro</td>
<td>0.45</td>
</tr>
<tr>
<td>Mozambique</td>
<td>0.89</td>
</tr>
<tr>
<td>Myanmar</td>
<td>0.32</td>
</tr>
<tr>
<td>Nepal</td>
<td>0.91</td>
</tr>
<tr>
<td>Niger</td>
<td>0.25</td>
</tr>
<tr>
<td>Nigeria</td>
<td>0.3</td>
</tr>
<tr>
<td>Pakistan</td>
<td>0.79</td>
</tr>
<tr>
<td>Peru</td>
<td>0.93</td>
</tr>
<tr>
<td>Philippines</td>
<td>0.87</td>
</tr>
<tr>
<td>Poland</td>
<td>0.34</td>
</tr>
<tr>
<td>Portugal</td>
<td>0.69</td>
</tr>
<tr>
<td>Rwanda</td>
<td>0.13</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>0.87</td>
</tr>
<tr>
<td>Senegal</td>
<td>0.94</td>
</tr>
<tr>
<td>Serbia</td>
<td>0.25</td>
</tr>
<tr>
<td>Sierra Leone</td>
<td>0.94</td>
</tr>
<tr>
<td>Slovenia</td>
<td>0.85</td>
</tr>
<tr>
<td>Somalia</td>
<td>0.95</td>
</tr>
<tr>
<td>South Africa</td>
<td>0.09</td>
</tr>
<tr>
<td>South Sudan</td>
<td>0.91</td>
</tr>
<tr>
<td>Sudan</td>
<td>0.17</td>
</tr>
<tr>
<td>Sweden</td>
<td>0.9</td>
</tr>
<tr>
<td>Switzerland</td>
<td>0.92</td>
</tr>
<tr>
<td>Syria</td>
<td>0.11</td>
</tr>
<tr>
<td>Tajikistan</td>
<td>0.92</td>
</tr>
<tr>
<td>Tanzania</td>
<td>0.09</td>
</tr>
<tr>
<td>Thailand</td>
<td>0.91</td>
</tr>
<tr>
<td>Togo</td>
<td>0.95</td>
</tr>
<tr>
<td>Tunisia</td>
<td>0.83</td>
</tr>
<tr>
<td>Turkey</td>
<td>0.51</td>
</tr>
<tr>
<td>Uganda</td>
<td>0.57</td>
</tr>
<tr>
<td>Ukraine</td>
<td>0.72</td>
</tr>
<tr>
<td>United Arab Emirates</td>
<td>0.91</td>
</tr>
<tr>
<td>USA</td>
<td>0.67</td>
</tr>
<tr>
<td>Yemen</td>
<td>0.07</td>
</tr>
<tr>
<td>Zambia</td>
<td>0.06</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Figure 3.21: Predicted Probabilities from MC LR Model
Appendix A
<table>
<thead>
<tr>
<th>Group</th>
<th>Factor</th>
<th>Discussed</th>
<th>Modeled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demand characteristics</td>
<td>Budget constraints</td>
<td>Salmerón and Apte, 2010; Mete and Zabinsky, 2010; Duran et al., 2011; Martel et al., 2013; Rennemo et al., 2014</td>
<td>Salmerón and Apte, 2010; Mete and Zabinsky, 2010; Duran et al., 2011; Martel et al., 2013; Rennemo et al., 2014</td>
</tr>
<tr>
<td></td>
<td>Demand risk</td>
<td>Ja et al., 2007; Salmerón and Apte, 2010; Mete and Zabinsky, 2010; Rawhi and Tunnquist, 2010; Duran et al., 2011; Gómez et al., 2011; Bozorgi-Amiri et al., 2013; Martel et al., 2013; Rennemo et al., 2014</td>
<td>Ja et al., 2007; Salmerón and Apte, 2010; Mete and Zabinsky, 2010; Rawhi and Tunnquist, 2010; Duran et al., 2011; Gómez et al., 2011; Bozorgi-Amiri et al., 2013; Martel et al., 2013; Rennemo et al., 2014</td>
</tr>
<tr>
<td>Logistics</td>
<td>Fleet size</td>
<td>Salmerón and Apte, 2010; Bozorgi-Amiri et al., 2013; Rennemo et al., 2014</td>
<td>Salmerón and Apte, 2010</td>
</tr>
<tr>
<td></td>
<td>Limitation in and/or possible damage to infrastructure</td>
<td>Ja et al., 2007; Salmerón and Apte, 2010; Mete and Zabinsky, 2010; Rawhi and Tunnquist, 2010; Duran et al., 2011; Gómez et al., 2011; Bozorgi-Amiri et al., 2013; Gahmada and Batta, 2013; Liberatore et al., 2013; Martel et al., 2013; Rennemo et al., 2014</td>
<td>Ja et al., 2007; Rawhi and Tunnquist, 2010; Salmerón and Apte, 2010; Gahmada and Batta, 2013; Liberatore et al., 2013; Martel et al., 2013; Rennemo et al., 2014</td>
</tr>
<tr>
<td></td>
<td>Supplier proximity</td>
<td>Duran et al., 2011; Gómez et al., 2011; Martel et al., 2013; Salmerón and Apte, 2010; Bozorgi-Amiri et al., 2013; Rennemo et al., 2014</td>
<td>–</td>
</tr>
<tr>
<td>Political and security situation</td>
<td>Personnel availability</td>
<td>Salmerón and Apte, 2010</td>
<td>Salmerón and Apte, 2010</td>
</tr>
<tr>
<td></td>
<td>Exchange rates, tariffs, tax incentives, customs clearing</td>
<td>Duran et al., 2011</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Level of unrest</td>
<td>Martel et al., 2013</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Long-term agreements with governments</td>
<td>Martel et al., 2013</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Price increases in the event of a disaster</td>
<td>Rawhi and Tunnquist, 2010; Bozorgi-Amiri et al., 2013; Gahmada and Batta, 2013; Liberatore et al., 2013</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Sociopolitical factors</td>
<td>Charni and Novo, 2008; Duran et al., 2011; Martel et al., 2013</td>
<td>–</td>
</tr>
</tbody>
</table>

Figure A-1: Overview of influencing factors identified in humanitarian network design literature.
Appendix B
<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Calculation/Assumptions</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transportation cost to satisfy ER demand by express shipment (see 4.3)</td>
<td>Air transportation: 100 USD per mile * total miles + 25,000 USD (fixed cost) / (divided by) value per TEU (twenty-foot container) Road transportation: 10 USD per mile * total miles / (divided by) value per TEU The first 10 TEUs in each new emergency event are shipped by air. Air shipments to any particular demand point are constrained by the availability of funding and cannot exceed 10 TEUs. Additional TEUs are sent via surface</td>
<td>Brambles, ongoing consultancy project at UNHCR Translated via analysis of ERP data Distances were calculated using the following link: <a href="http://www.freemaptools.com/how-far-is-it-between.htm">http://www.freemaptools.com/how-far-is-it-between.htm</a> For value per TEU, see below</td>
<td></td>
</tr>
<tr>
<td>Transportation cost to satisfy EO demand by normal shipment (see 4.3)</td>
<td>10 USD per mile * total miles / (divided by) value per TEU 20% of annual product value</td>
<td>Brambles, ongoing consultancy project at UNHCR Translated via analysis of ERP data Currently UNHCR does not operate with holding cost. Even so, variable cost was, in agreement with UNHCR, set to 20% of annual product value to account for variable staff cost, and so on.</td>
<td></td>
</tr>
<tr>
<td>Warehousing variable cost: holding cost per unit</td>
<td>Staff: USD 480,000 per year; USD 360,000 (international staff) + USD 120,000 (local staff); Rent of land/facility: USD 250,000 per year Total: USD 730,000 per year Assume warehouse rental so no opening or closing cost</td>
<td>Average warehousing fixed costs calculated and confirmed by UNHCR</td>
<td></td>
</tr>
<tr>
<td>Transportation lead time by air (for emergencies only) from confirmed order (and payment) until delivery at nearest airport</td>
<td>First wave after new emergency 3 days for all distances below 1000 miles 4 days for all distances above 1000 miles Set max transport by air for a new emergency to 430,000 USD (value of 10 containers) per year, the rest by sea/land</td>
<td>Brambles, ongoing consultancy project at UNHCR Translated by analyzing ERP data</td>
<td></td>
</tr>
<tr>
<td>Transportation lead time by sea/land (normal shipment for EO and for later waves of ER)</td>
<td>Estimated lead time 1 day per 200 miles</td>
<td>Brambles, ongoing consultancy project at UNHCR Translated by analyzing ERP data</td>
<td></td>
</tr>
</tbody>
</table>

Figure B-1: Details of data sources, assumptions, and cost and lead time calculations
Bibliography


