Data Science for Improving Operations in Organizations that Serve Vulnerable Populations

A Dissertation
Submitted to the Faculty
of the
WORCESTER POLYTECHNIC INSTITUTE
in partial fulfillment of the requirements for the
degree of

Doctor of Philosophy

in

Data Science

by Geri Louise Dimas

April 2023

APPROVED:

Professor Andrew C. Trapp
Worcester Polytechnic Institute
Advisor

Professor Renata A. Konrad
Worcester Polytechnic Institute
Co-Advisor

Professor Daniel Reichman
Worcester Polytechnic Institute
Committee Member

Professor Kayse Lee Maass
Northeastern
External Committee Member

Abstract

Data Science for Social Good continues to gain attention in research and the media. Data Science and Analytics can be used in many ways to aid vulnerable sectors of our society. I contribute to this effort by using analytics to improve the operations of governmental agencies, non-governmental organizations (NGOs) and nonprofit organizations serving vulnerable populations in two important domains: anti-human trafficking and immigration.

In my first investigation, I conducted an in-depth review of the current research landscape of Analytics and Operations Research as applied to the domain of anti-human trafficking. This review provides analysis for understanding, illuminating gaps, and proposing ways forward for those working at the intersection of Applied Analytics and Operations Research to fight human trafficking. In addition, I created a novel tool that allows researchers to explore further the meta-data associated with my review. In my next investigation, I examined the operational efficiency of an NGO working to fight human trafficking. I provided insights to aid this resource-constrained organization in improving its novel transit-monitoring operations at the Nepal / India border. In my third investigation, I studied the United States growing immigration court backlog by using a queueing-theoretic approach to examine the data, system structures, and behaviors. The derived framework provided an initial understanding of the United States Immigration court system. It was then used to develop a discrete event simulation model of the New York City immigration court system. Through this model, I demonstrated how changes in arrival and service rates affect key performance indicators (KPIs) of the system.

The developed model captures the complexity and interdependencies of the immigration court system and provides a foundation for further evaluation. Motivated by these insights, I extended the discrete event simulation model to enable an in-depth and robust exploration of how policy can impact and improve outcomes. In particular, I incorporated three policies supported by domain experts and evaluated the influence each has on reducing the KPIs of sojourn times, wait times, and queue lengths. The first policy varies the total number of judges and illustrates the impact the quantity of judges has on system throughput. The second policy prioritized asylum cases and evaluated the equity of dedicated dockets. The final policy aimed to minimize sojourn and wait times due to court-caused delays through the introduction of "make-up" capacity. The testing of such policies within my model provided initial data-informed insights for decision makers, something in critically short supply in this important aspect of our society. While the results demonstrate the power of using analytics to examine the immigration court backlog, collaborations with domain experts and other stakeholders are required to ensure a socially conscientious and reliable support tool for decision makers. The expanded model can then be used to test additional policies and seek more equitable solutions to better serve those in the immigration court system.

Acknowledgments

I would like to thank my advisor Dr. Andrew Trapp and co-advisor, Dr. Renata Konrad. Their guidance, support, and patience in both this work and my Ph.D. journey have been instrumental in reaching this milestone. I would like to thank my other committee members, Dr. Kayse Lee Maass and Dr. Daniel Reichman, for their contributions to this work.

I would like to thank all those in the Data Science community at WPI who have made the bad days good and the good days better – I am truly grateful to have been surrounded by such a wonderful and caring community. I would also like to thank Dr. Elke Rundensteiner, who has worked hard to cultivate the Data Science community at WPI as well as for all her kind words and support over the years. I thank Kelsey Briggs for the countless ways she has supported me.

I would also like to thank the Robert and Esther Goddard Fellowship for supporting the first year of my Ph.D. and the National Science Foundation (Operations Engineering grant CMMI-1825348 and CMMI-1935602) for supporting the remainder.

I want to thank Catalina and Brian– your presence and impact on my life are abundant and one of the reasons I ventured on this journey in the first place.

Nobody has been more important to me in the pursuit of this milestone than the members of my family. I thank my sisters for their support and always picking up the phone. I thank my nieces and nephews for showering me with unconditional love and giving me a reason to work hard. I thank my grandma for her love and encouragement over the years. I would like to thank my biggest supporters– my parents, who have always taught by example what it means to work hard and dream big. Their belief in my abilities and constant love, encouragement, and prayers have made all the difference. But above all else, I praise God for his faithfulness and all he has done in my life – Hallelujah, even here.

Finally, the completion of this Dissertation could not have been possible without the assistance and support of so many others whose names can not all be enumerated. I gratefully acknowledge and appreciate their contributions.

Contents

[1] Dimas, G. L., Konrad, R. A., Maass, K. L., Trapp, A. C. (2022). Operations research and analytics to combat human trafficking: A systematic review of academic literature. PLOS ONE 17(8): e0273708. https://doi.org/10.1371/journal.pone.0273708

[2] Dimas, G. L., El Khalkhali, M., Bender, A., Maass, K. L., Konrad, R., Blom, J. S., Zhu, J., Trapp, A. C. (2023). Estimating Effectiveness of Identifying Human Trafficking via Data Envelopment Analysis. Accepted, INFORMS Journal on Applied Analytics (Preprint available at https://arxiv.org/abs/2012.07746)

Introduction

The field of Data Science (DS) continues to develop, adapt, and be applied across all sectors of society. The growth and development of DS has been accompanied by a rise in interest and commitment to use these tools for social good. The DS for Social Good movement leverages DS methodologies to tackle a wide array of social and humanitarian problems such as algorithmic fairness [1], [2], refugee resettlement [3], [4], human trafficking [5], [6], [7], child welfare [8], and homelessness [9]. While the specific social motivation may differ, each application shares a common theme of utilizing the power of analytics to address an issue in society. Using DS to provide data-driven insights for improving operations and decision-making is well established and has conventionally been used in the private (for-profit) sector. However, the public sector, which includes nonprofit, governmental and non-governmental organizations (NGOs), provides immense and often overlooked opportunities to use DS to provide data-driven insights, thus the term "for social good." This dissertation contributes therein by using DS to improve the operations of government agencies, NGOs and nonprofit organizations serving vulnerable populations in two important domains: anti-human trafficking and immigration. Notably, I demonstrate the power and value of insights gleaned from data-driven methods for decision makers in these two contexts.

The organization of this report is as follows: Chapters 1 and 2 focus on my contributions to the anti-human trafficking domain. Chapter 1 provides an in-depth review of the current research landscape of Operations Research and Applied Analytics within the domain of anti-human trafficking. This review provides analysis on understanding what work is currently being done, illuminating gaps, and proposing ways forward for those working at the intersection of Operations Research and Analytics to address human trafficking. Notably, I introduce a novel tool that allows researchers to explore further the meta-data associated with the review for additional insights. Chapter 2 contributes to the growing literature examined in Chapter 1 through analyzing existing data and evaluating the performance of border stations of NGO Love Justice International (LJI) engaged in the anti-trafficking strategy known as transit-monitoring at the Nepal/India border. Through understanding the performance of these border stations, I provide operational improvement recommendations for LJI's decision makers.

Chapters 3 and 4 study the United States (U.S.) Immigration court system, seeking improvements to this complex system that aid in the reduction of the backlog and decrease average wait time for cases. Chapter 3 explores modeling the immigration court system via queuing theory. We utilize queuing theory to construct a mathematical representation of the United States immigration court system and capture the different states and processes within. Building off the mathematical representation and design in Chapter 3, Chapter 4 introduces a discrete event simulation model to simulate the NYC immigration court system from 2010-2019. The development of this baseline model provides insights into how changes in arrival and service rates affect key performance indicators of the NYC court system from 2010-2019.

Finally, in Chapter 5 we extend the model in Chapter 4 to enable a more in-depth and robust exploration of how policy can impact and improve outcomes. Three policies supported by our analysis in Chapter 4 and domain experts are incorporated and evaluated seeking to understand the effects each has on case sojourn times, wait times, and queue lengths. The first policy varies the total number of judges and illustrates the impact the quantity of judges has on system throughput. The second policy prioritized asylum cases and evaluated the equity of dedicated dockets. The final policy aimed to minimize sojourn and wait times due to court-caused delays through the introduction of "make-up" capacity. The testing of such policies can provide data-informed insights for decision makers, something in critically short supply in this important aspect of our society. The expanded model can further be deployed to test additional policies and can serve as a reliable decision-making tool for stakeholders to seek more efficient and equitable solutions to better support and serve those in the immigration court system.

Data Science for Social Good has profound potential to transform the world around us. The same tools and techniques that have transformed industry for decades are imminently available and, when attuned to the needs of society around us, have a powerful potential to effect change. This dissertation seeks to demonstrate tangible ways of using Data Science for Social Good to bring positive change to the domains of anti-human trafficking and immigration.

Chapter 1

Operations Research and Analytics to Combat Human Trafficking: A Systematic Review of Academic Literature [1]

Section 1.1 Introduction

Human trafficking (HT) involves the commercial exchange and exploitation of individuals for monetary or other gain using force, fraud, or coercion [10] and is a widespread social, economic, and human rights issue. While the trafficking of individuals is a centuries-old phenomenon, over the past two decades there has been growing public and research awareness, in part with the ratification of the 2000 Palermo Protocol to Prevent, Suppress, and Punish Trafficking in Persons [11]. Although precise figures are elusive, the Global Estimates of Modern Slavery Report estimates that HT impacts 25M individuals and annually generates more than 150 billion USD in illicit gains globally [12,13]. HT is broadly classified as labor and sex trafficking; while all trafficking features exploitation, the actions and means by which HT occurs may differ [14]. Labor trafficking takes place in a wide variety of sectors, including the agriculture, domestic work, construction, fishing, food service, and beauty industries. Sex trafficking is a part of the broader commercial sex industry, occurring in industries such as escort services, brothels, and pornography.

Because the scope of HT activity is vast and there are diverse ways in which individuals are exploited [15], context is critical, and effectively addressing HT increasingly requires efforts from multiple disciplines, including

[1] Dimas, G. L., Konrad, R. A., Maass, K. L., Trapp, A. C. (2022). Operations research and analytics to combat human trafficking: A systematic review of academic literature. PLOS ONE 17(8): e0273708. https://doi.org/10.1371/journal.pone.0273708

interdisciplinary collaborations. For example, HT interventions include approaches from multiple sectors and disciplines such as social work [16–18], healthcare [19, 20], criminal justice [21–23], and economics [24, 25]; each domain brings unique perspectives and methods to understand and address HT.

Owing to the breadth of domains that contribute to anti-HT research, a wealth of literature exists that has been well-documented in surveys over the years [19,20,26–30]. Existing reviews focus on social science, healthcare, and law enforcement approaches; whereas OR and Analytics have much to offer [31], no systematic review exists for the emerging landscapes of Operations Research (OR) and Analytics as applied to anti-HT.

The present study identifies and classifies the existing OR and Analytics literature related to anti-HT operations. Building off the earlier work of Krammer-Kerwick et al. [32] and Caulkins et al. [33], this systematic review proposes an agenda for future research in this field, filling a gap in the current literature. This study focuses on the four broad principles of anti-HT: prevention, protection, prosecution, and partnership (4Ps) [34], extending their definition in relation to the OR and Analytics fields. We examine the following research questions:

(i) What aspects of HT are being studied by OR and Analytics researchers?

(ii) What OR and Analytics methods are being applied in the anti-HT domain?

(iii) What are the existing research gaps associated with (i) and (ii)?

We organize the remainder of our study as follows. In the method of collection and categorization section we define the scope of this review, and in the data section we define the data features for analysis. In the implications and observations section we discuss the implications of the survey and, based on the observed gaps, suggest areas for future work. We conclude our study in the final section.

Section 1.2 Method of Collection and Categorization

We conducted a systematic literature review inventorying studies to answer the three research questions outlined in the introduction. The methodology used for this systematic review was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [35]. The collection process (Fig 1.1) was based on keyword searches that generated a set of research studies for analysis. Two sets of search words were defined using the combined knowledge of the authors on HT terminology and OR and Analytics methods. These keywords were used in a procedure to identify and select studies that met a set of pre-defined criteria. The first set of keywords reflects terms related to HT, while the second reflects common methods in the OR and Analytics fields (see Table 1.1). The search and selection of studies was performed by the lead author (G.L.D.), and any uncertainty regarding a study's inclusion was resolved through discussion with the coauthors.

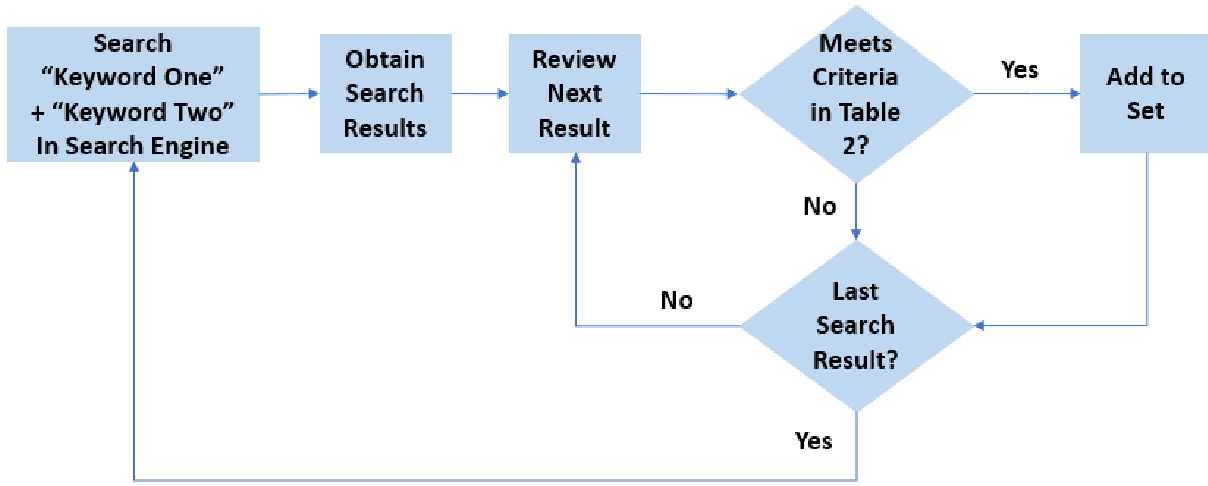

Figure 1.1: **Heuristic Process of Data Collection.**

Each search query followed the format: "Keyword One" + "Keyword Two" (such as "Debt Bondage" AND "Integer Programming"), each keyword pair was applied across three bibliographic databases: Scopus, Web of Science, and Google Scholar. The database search was conducted from June 2021 through March 2022. The search results were truncated to studies available through the end of 2021 to provide a comparable year-over-year basis for the research landscape. The sum of these two searches resulted in a total of 449,407 studies for potential inclusion. After the keyword search identification process, a two-step selection process was followed (see Fig 1.2). An initial screening process was conducted that evaluated the search results returned for each query, where titles and abstracts were screened and added to the set based on the criteria outlined in Table 1.2. The stopping criteria for each keyword pair search followed a heuristic approach: if at least 50 results returned no eligible results, then the current keyword search was stopped and the next keyword search began. The intuition behind this approach is that by design, search engines return the most relevant results first, and therefore if after a certain point no relevant results are produced (at least 50 results in our context), then it is highly unlikely relevant results exist past that point. After the initial screening process, the set included 230 unique studies for review. A more in-depth review using the eligibility requirements checklist (Table 1.2) was followed in step two for each of the 230 studies.

First, only studies that fell into one of the three themes: *Operations Research* methodologies, *Analytical* methodologies, or *Position / Thought* pieces related to Operations Research and Analytics were included. Second, only studies whose primary application area was anti-HT and was written completely in English text were included. Third, book chapters, workshops, and reports (governmental, intragovernmental, non-government organizations,

Table 1.1: **Keywords Used in Search Process.**

Keyword One: Human Trafficking Related	
Child Labor / Child Labour	Debt Bondage
Domestic Servitude	Forced Labor/ Forced Labour
Human Trafficking	Labor Trafficking / Labour Trafficking
Modern Slavery	Sex Trafficking
Trafficking in Persons	
Keyword Two: Methodology Related	
Clustering/Classification	Data Envelopment Analysis
Data Science	Game Theory
Graph Theory/Construction	Information Extraction
Integer Programming	Machine Learning
Natural Language Processing	Network Interdiction/Flow
Operations Research	Queueing/Queueing Theory
Resource Allocation	Simulation
Simulation	Supervised/Unsupervised Learning
Supply Chain	Web Crawling

think-tanks) were excluded. Finally, if multiple studies related to a single work were found, only the most complete version were included. Peer-reviewed studies, dissertations, and pre-print studies were included to produce a full and comprehensive review of the current research landscape. The full article screening process resulted in a total of 142 studies included in the set for final review.

As with any search process based upon a predefined set of keywords and checklist requirements, the 142 identified studies may not be exhaustive in scope. However, given our collective experience in researching at the intersection of anti-HT and OR and Analytics, we believe the generated set of studies is representative of current literature at this confluence. A repository containing our classification data can be found publicly at `https://github.com/gldimas/Dimas-et.-al-2022_Human-Trafficking-Literature-Review`.

While the focus of this study was to capture the general trends in anti-HT research within the OR and Applied Analytics communities, we acknowledge there exist many relevant studies that either fall outside of the scope of the present analysis (such as reports) or focus on the problem domain of anti-HT but do not have a specific OR or Analytics-related attribution, and therefore were excluded. We have included such studies in the S2 File and

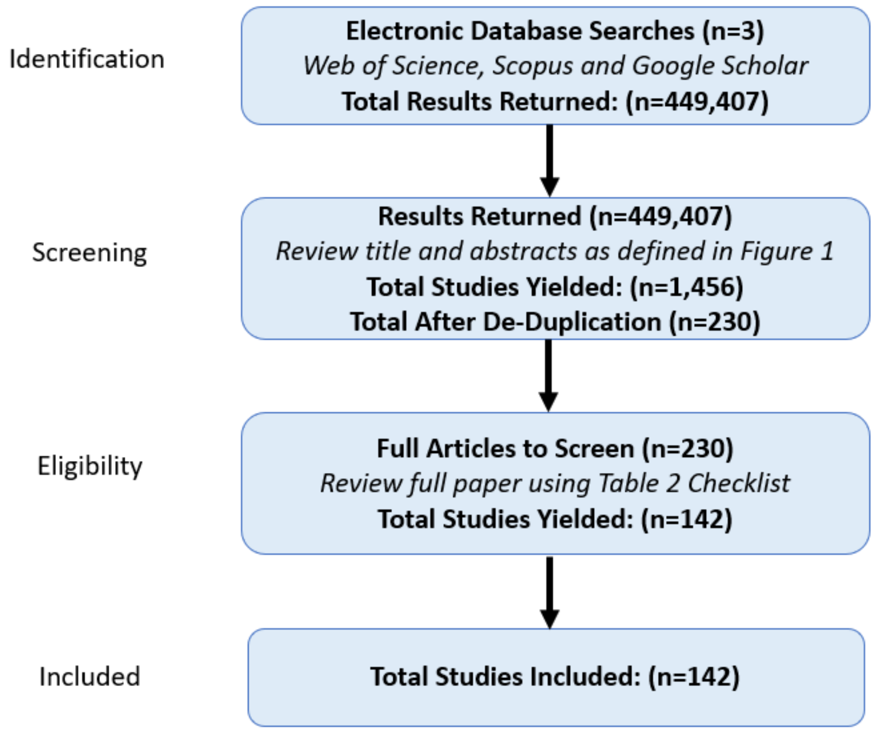

Figure 1.2: **Overview of Selection Process.**

Table 1.2: **Eligibility Requirements Checklist.**

Requirements Checklist
1. Main contribution or focus fell into one of the following three themes:
• Methodological Operations Research orientation
• Methodological Analytics orientation (Data Science, or other Applied Analytics)
• Position / Thought pieces in Operations Research or Analytics
2. Main application or case study was on anti-HT efforts
3. Only studies, articles, theses and dissertations were kept (e.g., no books, workshops or government reports)
4. If multiple versions of a study exist (such as a conference paper followed by a peer-reviewed journal article) only the most recent, comprehensive version was kept
5. If the study appeared as a section in one study but was further developed into a full study, only the full study was kept

further point the reader to reviews such as Raets and Janssens [36], Farrell and De Vries [37], and Weitzer [38] to better capture the broader scope of anti-HT research.

Section 1.3

Data

The classification of studies was independently conducted by the lead author (G.L.D.) who throughout the process conferred with all coauthors. Each of the 142 studies in the set were reviewed and assigned labels to nine key features: **Publication Year**, **Category**, **Context**, **Demographics**, **Target Region**, **Data Source**, **Theoretical Approach**, **Methodologies**, and **4Ps**. We next explain each feature and its respective values.

Publication Year: Observed years were **2010 − 2021**, inclusive. Only studies available through December 2021 were included in our scope to allow for comparison across complete calendar years. This feature offers valuable information about the progress and patterns of research in OR and Analytics over more than ten years.

Category: We considered three categories: **Operations Research (OR)**, **Analytics**, and **Position / Thought**. Whereas the first two categories are distinguished by their *methodological* focus, the latter includes *position / thought* pieces from either the OR or Analytics domains. We required a single category to be assigned to each study, and thus selected the category that we felt best matched the primary theme of the work.

Context: We classified the primary topical HT area (as stated or inferred) into three contexts: **Sex**, **Labor**, and **General**. If the application was not specified, we assumed it to be general and thus applicable to both sex and labor trafficking.

Demographics: We classified studies into five demographic groups based on the population of interest (such as victims, potential victims, and survivors): **Female**, **Male**, **Child**, **LGBTQ+**, and **Unspecified / All Individuals**. If no specific demographic characteristics were stated or could be inferred, we assumed it to be applicable to all individuals. A study could be classified into multiple demographics such as a study focused on female children.

Target Region: We subdivided the geographic location specified either by the data used in the study or by the region discussed in the background of the study, into world regions: **Africa**, **Asia**, **Australia/Oceania**, **Europe**, **North America**, **South America**, **Unspecified / All regions**. A study could cover multiple geographic locations and therefore have multiple target regions identified.

Data Source: We classified the type of data used in the study into four categories: **Primary**, **Secondary**, **Mixed**, **N/A**. Primary data are collected directly from anti-HT organizations or researchers, including interviews and surveys. Secondary data are data that have already been collected for other purposes or are publicly available such as data from websites hosting illicit advertisements (such as backpage.com and rubmaps.com) and government reports. While many studies used their own methodologies to scrape public data sources such as escort and massage websites, we still consider these to be secondary sources. We classify studies utilizing expert judgements for determining data estimates to be secondary data. Mixed data means the study used both primary and secondary data in their work, and N/A indicates data was not used in the study.

Theoretical Approach: We classified the central theoretical approach to address HT into six categories: **Decision Support**, **Inferential Statistics / Detection**, **Network Flow**, **Resource Allocation**, **Supply Chain**, and **Other / Unspecified**. Decision Support explores ways to inform decision makers about initiatives to improve and better address HT, often building tools or systems for practitioners to use. Inferential Statistics / Detection focuses on identifying, estimating, or inferring aspects of HT. Network Flow studies are related to the flow of individuals and possibly trafficking network interaction. Resource Allocation addresses the use and allocation of resources in anti-HT efforts. Supply Chain studies examine the supply and demand of HT within a network. If a study approaches HT from a theoretical approach not listed, we label these studies as Other / Unspecified. A study may be classified under multiple theoretical approaches.

Methodologies: We classified the main methodologies used in the set of studies into 21 categories:

- Active Learning
- Clustering or Classification
- Data Envelopment Analysis
- Empirical Analysis
- Facility Location
- Game Theory
- Graph Construction
- Investigative Search
- Link Inference
- Machine / Deep Learning (General)
- Natural Language Processing
- Information Extraction
- Integer Programming
- Network / Graph Theory
- Network Interdiction
- Queueing Theory
- (Social) Network Analysis
- Simulation
- Unsupervised or Minimally Supervised Learning
- Web Crawling / Scraping
- Other

We ascribe methods to a study based on the introduction, conclusion, and main method or focus throughout the study. A study may apply a variety of different methods and therefore be classified into multiple methodologies.

4Ps: Activities to fight HT are often discussed under four broad principles: prevention, protection, prosecution, and partnership. These principles are collectively referred to as the 4Ps – a well-recognized classification within the anti-HT community [31]. A study may be classified under multiple principles. The 4Ps naturally correspond with efforts in the social science, healthcare, and law enforcement disciplines, and their alignment with OR and Analytics works is less evident. Thus, we adapt the 4Ps definitions to define each as it relates to OR and Analytics using the collective knowledge and experience of the authors in the anti-HT and OR and Analytics fields. To the best of our knowledge, this is the first attempt to define each of the 4Ps as it relates specifically to the OR and Analytics fields and constitutes an important contribution of this study. Prevention, Protection, and Prosecution were originally referred to as the 3P paradigm [34] which has since been informally expanded to include a fourth "P" representing Partnership. Prevention refers to efforts focused on a proactive approach to prevent trafficking such as awareness campaigns and education; Prosecution refers to efforts to punish traffickers; and Protection involves meeting post-trafficking victim needs such as counseling, job training, housing, and other support to facilitate survivor recovery and restoration.

Partnership was introduced to serve as a complementary means to further improve the efficacy among the 3Ps, enlisting all segments of society in the fight against HT [34]. Together the 4Ps capture the spectrum of efforts in combating HT and therefore are an important feature for our literature review. Accordingly, we have adapted these 4Ps and classified studies in the following manner:

- **Prevention:** The goal of the study is the prevention of HT either now or in the future and assumes no trafficking is currently taking place. Such studies typically feature victim-centric methodologies to help potential victims avoid being trafficked, such as awareness campaigns and education. Studies that consider reducing the re-trafficking risk of survivors who have left their trafficking environment also fall within the scope of the prevention principle.

- **Protection:** The goal of the study is to protect and aid the survivor during and post-exploitation. We consider victim-driven detection and disruption of HT networks to be a form of protection, as the focus of the study is mitigating the risk to an individual of further exploitation (including studies that consider NGOs, healthcare, and other non-law enforcement detection).

- **Prosecution:** The goal of the study is to aid the prosecution of traffickers (often from a law enforcement perspective). We consider detection and disruption of HT networks aimed at locating, understanding, and stopping traffickers under the prosecution principle.

- **Partnership:** The goal of the study is to show the benefit of collaboration and data sharing across different sectors, countries, or groups working together toward the common goal of addressing one or more areas of HT.

Section 1.4

Implications and Observations

Of the 142 studies in the set, the majority (73.9%) were categorized as Analytics [7,39–142], with 15.5% classified as Operations Research [143–164] and 10.6% as Position / Thought [5,31,33,165–176]. Fig 1.3 depicts this breakdown. Fig 1.4 provides summary statistics on each of the nine key features. All percentages are calculated in relation to the total number of studies in the set (142) unless stated otherwise. As some studies contained multiple methods or were identified to have multiple values within a feature, a feature may not always sum to 100.0%.

Research question 1: What aspects of HT are being studied?

Although both sex and labor trafficking have been addressed in the OR and Analytics literature, an overwhelming number of studies focus specifically on sex trafficking. Fig 1.4 illustrates the inclination of OR and Analytics studies to focus on sex trafficking (47.9%), with only 12.0% concentrating on labor trafficking, while 40.1% apply to general (both) trafficking contexts. As observed in Fig 1.5, studies overwhelming use secondary data, with fewer than 4.0% using a primary data source. The use of secondary data is likely due to accessibility; almost all studies on sex trafficking (60.3%) used data pulled from escort websites (or other online sites hosting illicit advertisements) which are public and therefore easier to access. The use of escort websites (in particular, backpage.com) as a source of data result in over 38.0% of the studies focusing on the North American region. Although the United States Department of Justice shut down backpage.com in 2018 [177], other escort and massaging sites offer illicit services and constitute the data source for several studies. Remarkably, 76.8% of studies were not tailored to a specific demographic, despite the differences between typologies and demographics of victims [15]. From the 4Ps perspective, prosecution is the most common principle (62.7%) among all studies, with considerably less focus on partnership, protection, and prevention (Fig 1.6). A single study may be categorized under multiple 4Ps principles, and therefore the values in Fig 1.6 do not sum to 100.0%.

Research question 2: What OR and Analytical methods are being applied in the anti-HT domain?

Fig 1.4 provides summary statistics such as counts and percentages for the frequency each value was observed in the set of studies. While the percentages are calculated based on the 142 studies in the set, because a study may belong to multiple values within a feature, the total observations within each feature are provided in the parenthesis. The colors highlight the magnitude of studies within each feature value, with green indicating higher percentages, and red lower percentages. A study may fall under more than one value and therefore the percentages will not always sum to 100.0%. Methods related to machine learning (Machine / Deep Learning (General),

Analytics		
Acharjee et al. (2018)	Ibanez and Suthers (2014)	Mordeson, Mordeson and Mathew (2018)
Adam et al. (2021)	Ibanez and Suthers (2016)	Mukherjee et al. (2021)
Akram and Zafar (2019)	Imperial (2021)	Nagpal et al. (2015)
Ali et al. (2021)	Kejriwal and Gu (2020)	Poelmans et al. (2012)
Alvari et al. (2016)	Kejriwal and Kapoor (2019)	Portnoff et al. (2017)
Alvari et al. (2017)	Kejriwal and Szekely (2017a)	Prashad et al. (2021)
Andrews et al. (2018)	Kejriwal and Szekely (2017b)	Price (2016)
Ardakani (2020)	Kejriwal and Szekely (2017c)	Progga et al. (2020)
Asif et al. (2021)	Kejriwal et al. (2017)	Rabbany et al. (2018)
Binu et al. (2019)	Kejriwal et al. (2018)	Rodrigues et al. (2015)
Boyd et al. (2021)	Kennedy (2012)	Sabon, Yang and Zhang (2021)
Burbano et al. (2018)	Kiss et al. (2021)	Samanta et al. (2020)
Caoli (2019)	Kulshrestha (2021)	Sebastian et al. (2020)
Chambers et al. (2019)	Lavelle-Hill et al. (2021)	Sethi et al. (2013)
Chen et al. (2015)	Lee et al. (2021)	Silva et al. (2014)
Cockbain et al. (2011)	Li et al. (2018)	Simonson (2021)
Cockbain et al. (2019)	Li et al. (2021)	Stylianou et al. (2017)
Coxen (2021)	Libaque-Saenz et al. (2018)	Stylianou et al. (2019)
da Silva Santos et al. (2019)	Liu et al. (2019)	Szakonyi et al. (2021)
Darabian and Borzooei (2018)	Malik et al. (2018)	Szekely et al. (2015)
de Vries and Radford (2021)	Mancuso (2014)	Tahir et al. (2021)
Diaz and Panangadan (2020)	Mathew and Mordeson (2017)	Thöni et al. (2018)
Diviák et al. (2021)	Mathew, Mordeson and Yang (2019)	Tong et al. (2017)
Dubrawski et al. (2015)	McDonald et al. (2021)	Upadhayay et al. (2020)
Edge et al. (2020)	Mensikova and Mattmann (2018)	Vogt (2016)
Esfahani et al. (2019)	Mletzko et al. (2018)	Wang et al. (2012)
Gakiza et al. (2021)	Mordeson and Mathew (2017a)	Wang et al. (2020)
Giacobe et al. (2016)	Mordeson and Mathew (2017b)	White, Guikema, and Carr (2021)
Giommoni and Ikwu (2021)	Mordeson and Mathew (2018)	Whitney (2017)
Goist et al. (2019)	Mordeson and Mathew (2020a)	Wiriyakun and Kurutach (2021a)
Granizo et al. (2020)	Mordeson and Mathew (2020b)	Wiriyakun and Kurutach (2021b)
Harmon et al. (2020)	Mordeson et al. (2017)	Yang et al. (2018)
Hernández-Álvarez (2019)	Mordeson et al. (2019)	Yao et al. (2021)
Hultgren et al. (2018)	Mordeson, Mathew and Acharjee (2018)	Zhou et al. (2016)
Ibanez and Gazan (2016)	Mordeson, Mathew and Binu (2019)	Zhu et al. (2019)
Position/Thought		
Amin (2010)	Deeb-Swihart et al. (2019)	LeBaron (2021)
Borrelli and Caltagirone (2020)	Hundman et al. (2018)	McKenzie (2019)
Brewster, Ingle, et al. (2014)	Khanal (2020)	Orantes (2018)
Brewster, Polovina, et al. (2014)	Konrad et al (2021)	Tambe and Tambay (2020)
Caulkins et al. (2019)	Konrad et al. (2017)	Weinberg et al. (2020)
Operations Research		
Bhaumik et al. (2020)	Kosmas et al. (2020)	Neely et al. (2019)
Brelsford and Parakh (2018)	Kosmas et al. (2021)	Ramchandani et al. (2021)
Dimas et al. (2021)	Kougkoulos et al. (2021)	Senft et al. (2019)
Gerry et al. (2021)	Kővári and Pruyt (2014)	Stapleton et al. (2012)
Grimes et al. (2011)	Maass et al. (2020)	Taylor (2018)
Kapoor et al. (2017)	Mahbub (2021)	Tezcan and Maass (2020)
Keskin et al. (2021)	Mahdiraji et al. (2020)	
Konrad (2019)	Meier and Vitor (2021)	

Figure 1.3: **Category Classification for the Set of 142 Studies.**

Clustering or Classification, Unsupervised / Minimally Supervised Learning, Natural Language Processing, and Active Learning) were observed in over half of the studies. Machine / Deep Learning (General) and Clustering or Classification were the two most popular methods, accounting for about 32.0% of all methods observed (out of the total 404 methods identified, see Fig 1.4). Web Crawling / Scraping was used to generate a secondary dataset

Trafficking Context (n=142)	Count	% Total	Category (n=142)	Count	% Total
Sex Trafficking	68	47.9%	Operations Research	22	15.5%
Labor Trafficking	17	12.0%	Analytics	105	73.9%
General	57	40.1%	Position / Thought	15	10.6%
4P's (n=261)	**Count**	**% Total**	**Data Source (n=142)**	**Count**	**% Total**
Prevention	47	33.1%	Primary	5	3.5%
Protection	64	45.1%	Secondary	111	78.2%
Prosecution	89	62.7%	Mixed	8	5.6%
Partnership	61	43.0%	N / A	18	12.7%
Target Region (n=149)	**Count**	**% Total**	**Demographics (n=146)**	**Count**	**% Total**
Africa	3	2.1%	Female	17	12.0%
Asia	18	12.7%	Male	0	0.0%
Australia / Oceania	2	1.4%	Child	20	14.1%
Europe	16	11.3%	LGBTQ+	0	0.0%
North America	54	38.0%	Unspecified / All individuals	109	76.8%
South America	5	3.5%	**Methodologies (n=404)**	**Count**	**% Total**
Unspecified / All regions	51	35.9%	Active Learning	2	1.4%
Theoretical Approach (n=202)	**Count**	**% Total**	Clustering or Classification	60	42.3%
Decision Support	65	45.8%	Data Envelopment Analysis	1	0.7%
Inferential Statistics / Detection	69	48.6%	Empirical Analysis	21	14.8%
Network flow	33	23.2%	Facility Location	1	0.7%
Resource Allocation	3	2.1%	Game Theory	3	2.1%
Supply Chain	10	7.0%	Graph Construction	31	21.8%
Other / Unspecified	22	15.5%	Information Extraction	16	11.3%
Publication Year (n=142)	**Count**	**% Total**	Integer Programming	4	2.8%
2010	1	0.7%	Investigative Search	4	2.8%
2011	2	1.4%	Link Inference	1	0.7%
2012	4	2.8%	Machine / Deep Learning (General)	69	48.6%
2013	1	0.7%	Natural Language Processing	31	21.8%
2014	6	4.2%	Network / Graph Theory	17	12.0%
2015	5	3.5%	Network Interdiction	4	2.8%
2016	7	4.9%	Queueing Theory	0	0.0%
2017	15	10.6%	(Social) Network Analysis	24	16.9%
2018	21	14.8%	Simulation	11	7.7%
2019	22	15.5%	Unsupervised or Minimally Supervised Learning	19	13.4%
2020	22	15.5%	Web Crawling / Scraping	26	18.3%
2021	36	25.4%	Other	60	42.3%

Figure 1.4: **Summary Statistics for the Set of 142 Studies.**

for analysis in 18.3% of the studies, reflecting our previous observations in Research question 1 regarding the high use of secondary data extracted from massage and escort websites. At least one method for the construction and analysis of networks (Graph Construction, Network / Graph Theory, Network Interdiction and (Social) Network Analysis) were observed in 34.5% of the studies, with most emphasizing Graph Construction.

The observed theoretical approaches are closely related to specific methods. For example, nearly half of the studies focused on Inferential Statistics / Detection or Decision Support, most of which applied various machine

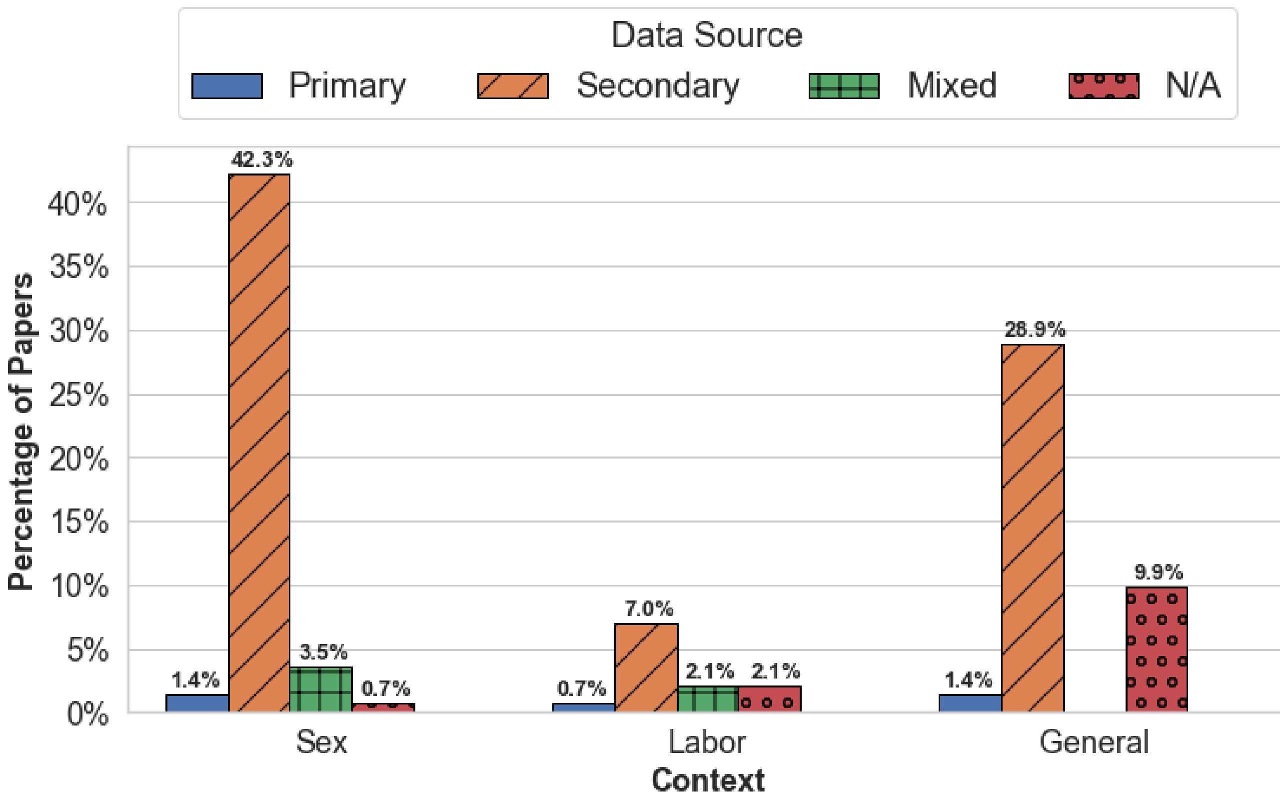

Figure 1.5: **Context of HT and Data Source.**

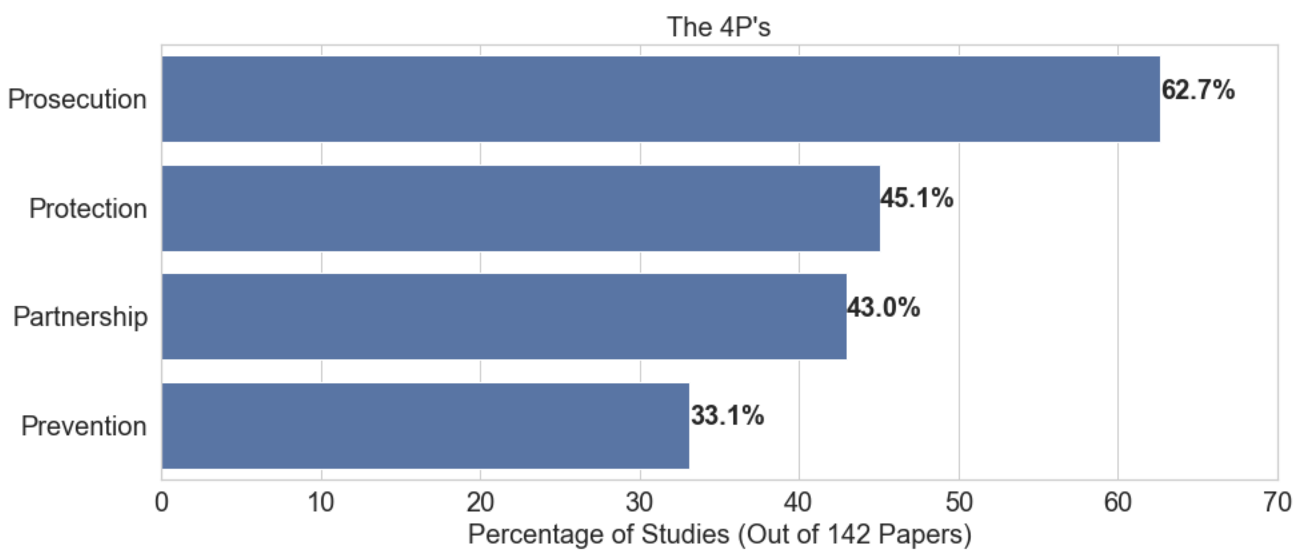

Figure 1.6: **Percent of Studies Involving 4Ps.**

learning methods. Network Flow methods appear in nearly 24.0% of studies, specifically Graph Construction and (Social) Network Analysis.

Fig 1.7 and 1.8 depict each study on the x-axis. In Fig 1.7 we display all studies categorized as Analytics, and in Fig 1.8 we categorize studies on the left as Operations Research (in blue), and studies on the right as Position /

Thought studies (in red). For each study, the top six y-axis labels indicate its Theoretical Approach(es), while the remaining 21 following labels indicate methods used. If a study includes a given feature value, the box is black, and grey otherwise. Theoretical approaches and methods are sorted in descending order based on the total count for each row. Fig 1.7 indicates that the majority of studies (over 77.0%) in the Analytics category take an Inferential Statistics / Detection or Decision Support theoretical approach. Fig 1.8 shows that studies in the Operations Research category are diverse, addressing the problem from distinct theoretical approaches and applying a variety of methodologies. Studies categorized as Position / Thought address a variety of theoretical approaches and topics, which is a good indication that OR and Analytics researchers are exploring HT from different fields.

While Resource Allocation and Supply Chain roughly make up only 9.0% of all studies, they account for over 40.0% within the Operations Research category. Looking more closely at the relationship between Web Crawling / Scraping and Clustering and Classification methods we see a large majority (nearly 70.0%) of Web Crawling / Scraping studies apply Clustering and Classification methods. In addition, more than 72.0% of studies that applied both Web Crawling / Scraping and Clustering and Classification shared the goal of identifying sex trafficking in online advertisements or tweets.

To compare studies across all three categories: Analytics, Operations Research, and Position / Thought Fig 1.7 and 1.8 are combined and the resulting figure can be found in the supplementary materials (see S1 Fig). We also provide a publicly available spreadsheet that can be used for closer examination of the individual studies, at `https://github.com/gldimas/Dimas-et.-al-2022_Human-Trafficking-Literature-Review`. This spreadsheet allows filtering studies on any combination of our nine features, returning all qualified studies. A screenshot of this tool can be see in Fig 1.9. As seen in this review, research at the intersection of anti-HT and OR and Applied Analytics is rapidly evolving and there is necessary complexity in the reviewing process. The selected keywords, search engines used and growing literature streams may impact the articles included in this review. In an effort to counter this limitation and support the longevity of this study, the authors have created an online submission form where individuals can submit works they believe to be related to the present study. Related submissions will be added to the dashboard tool on a semi-regular basis (see Fig 1.9) and will provide an evolving source of knowledge for researchers. The link for this form can be found on the github link provided earlier in this paragraph.

Research question 3: What are the existing gaps and opportunities for future research?

Several gaps emerged based on the classification and grouping of all 142 studies in the set on the prime principles (4Ps) (Fig 1.6), methods (Fig 1.7, and 1.8), trafficking context (Fig 1.5), and resulting observations to Research questions 1 and 2. The following sections present opportunities for new avenues of investigation for OR and

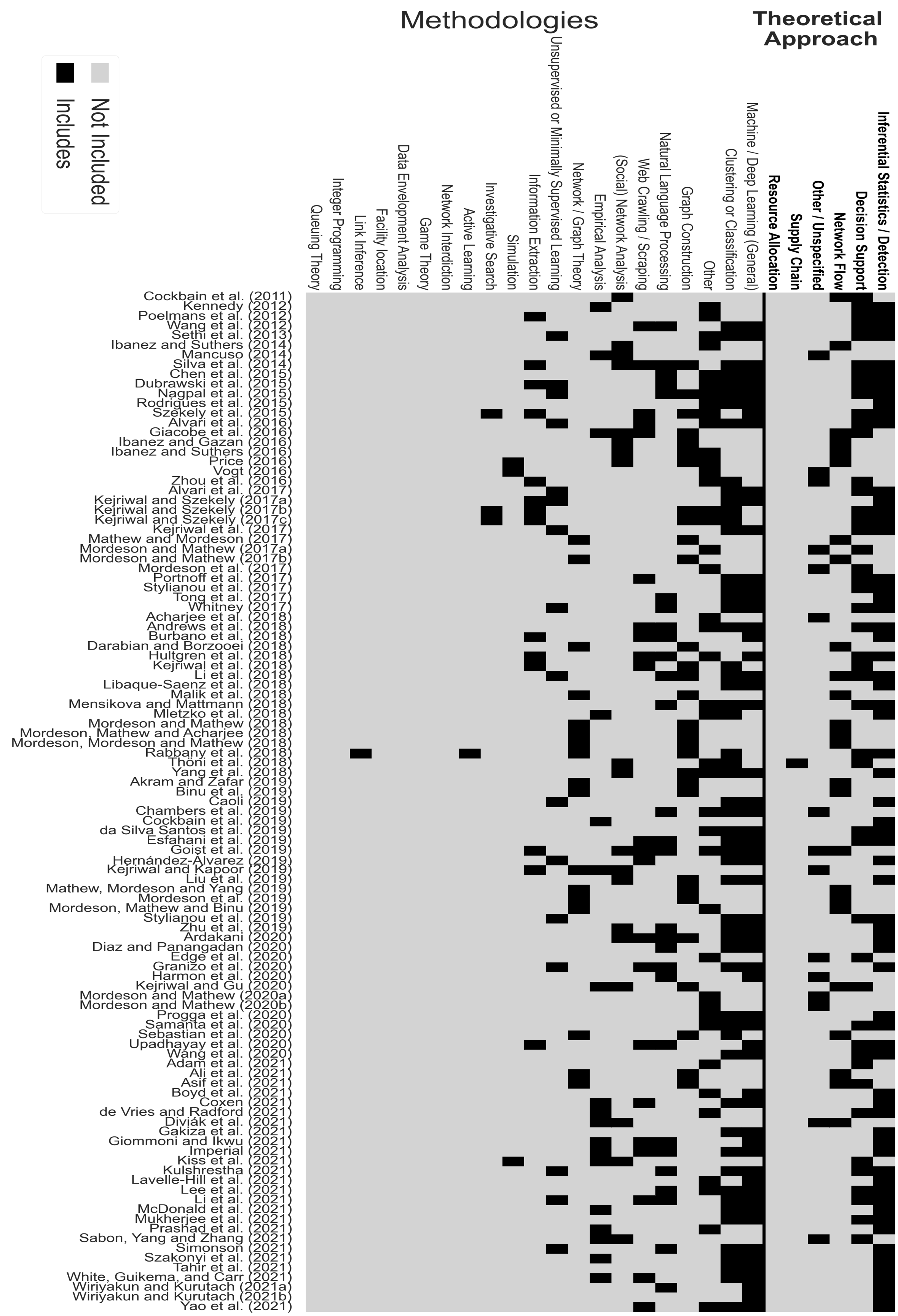

Figure 1.7: **Granular View of Theoretical and Methodological Topic Inclusion for the Set of Studies Categorized as Analytics.** The x-axis lists each study; and the y-axis depicts each of the Theoretical Approaches and Method. If a study includes a given feature value, the box is black, and grey otherwise. Theoretical approaches and methods are sorted in descending order based on the total count for each row.

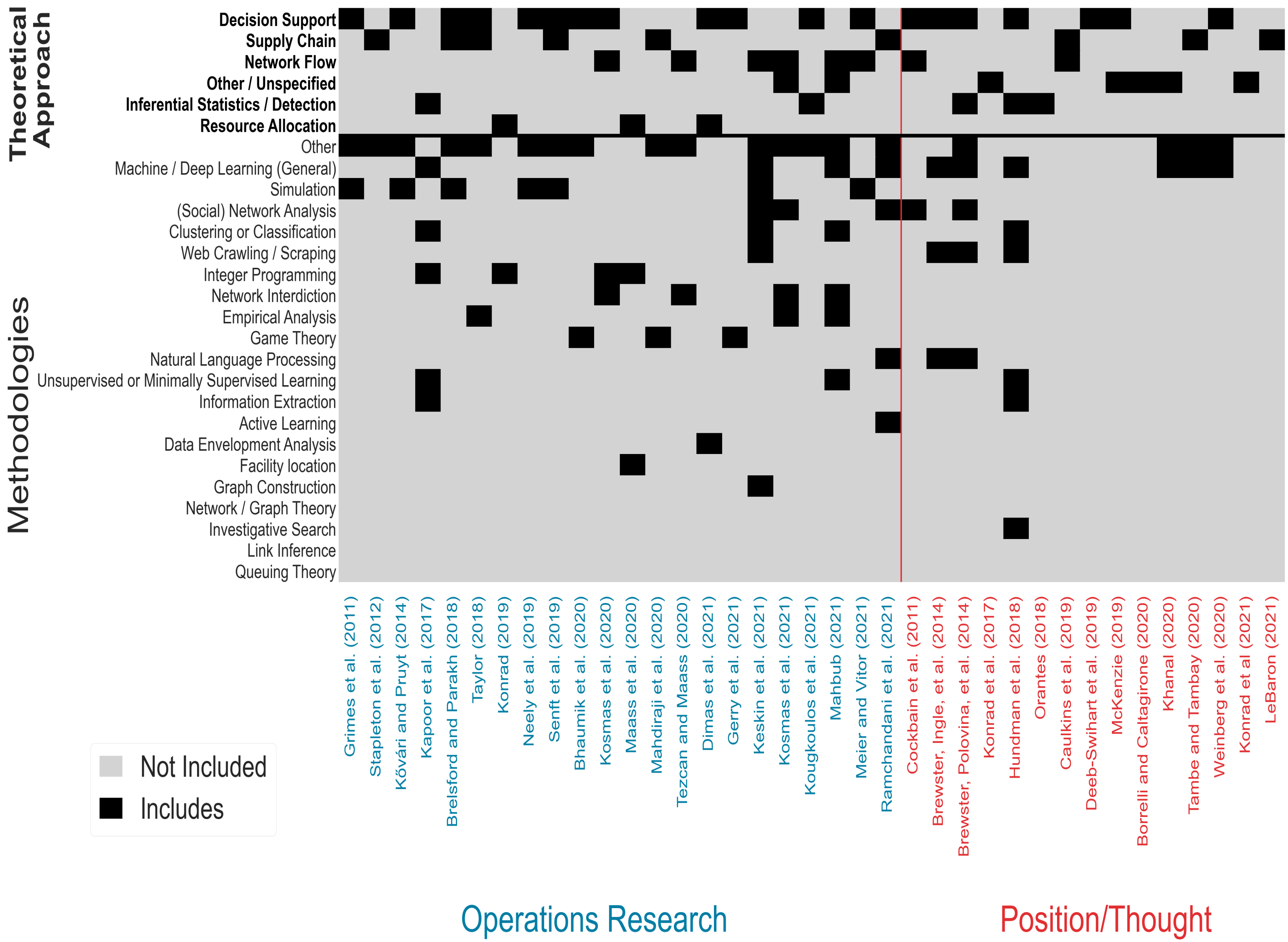

Figure 1.8: **Granular View of Theoretical and Methodological Topic Inclusion for the Set of Studies Categorized as Operations Research or Position / Thought**

The x-axis lists each study; and the y-axis depicts each of the Theoretical Approaches and Method. Operations Research studies appear on the left (in blue), and Position / Thought studies appear on the right (in red). If a study includes a given feature value, the box is black, and grey otherwise. Theoretical approaches and methods are sorted in descending order based on the total count for each row.

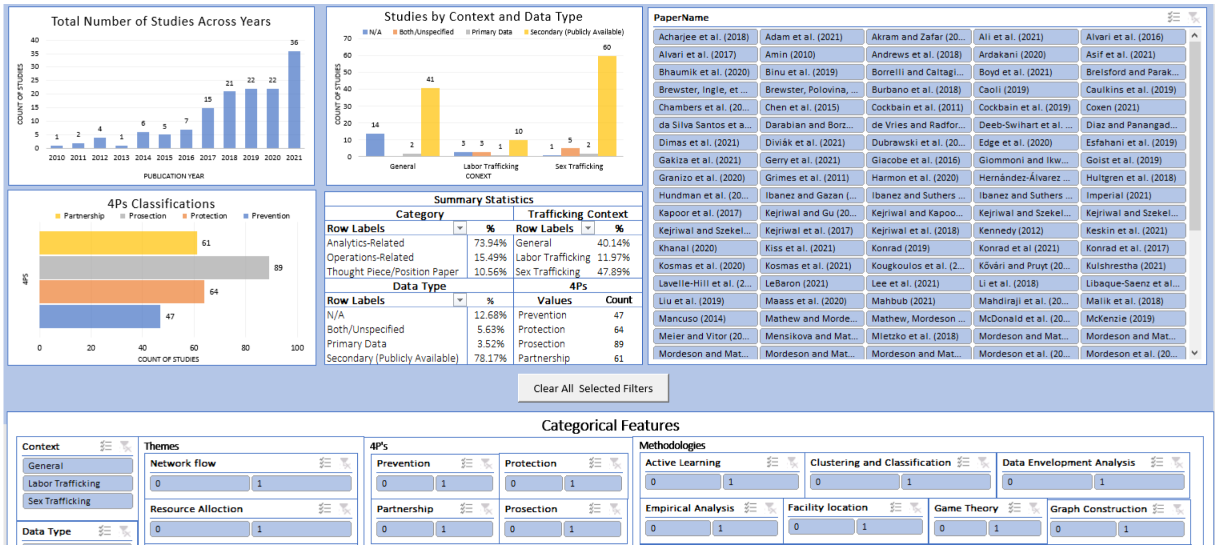

Figure 1.9: **A Screenshot of The Spreadsheet Tool Created For Closer Examination of the Set of 142 Studies.**

Analytics researchers in anti-HT efforts.

Broaden the typology and demographics of trafficking studied. The typology of trafficking activity, and by extension the demographic composition of victims, is diverse and under-explored in OR and Analytics [15]. As noted in [5], OR and Analytics researchers can increase the relevance and impact of their work by understanding the typology of trafficking and distinguish between various trafficking business models. To start, there is a clear need to expand the current focus to include labor trafficking. In OR and Analytics, sex trafficking (Fig 1.6) has received by far the most attention; while labor trafficking is estimated to account for over 60.0% of all trafficking instances [12], it constitutes only 12.0% of the reviewed studies. There is a close relationship between data source and type of HT studied. The absence of reliable, available data likely contributes to the lack of both analytically-based labor trafficking research, as well as the lack of diversity in data used in sex trafficking research; this is a well-documented issue across the anti-HT literature [31, 178–181]. The implications of the lack of data are also evident in studies using OR and Analytics methods and discussed further in the next section.

The diverse populations of those experiencing trafficking warrants further analysis. Nearly 77.0% of the studies observed were classified as applicable to Unspecified / All Individuals, indicating that the nuanced differences in victimology may be lacking. For example, no study we observed looked at trafficking through the lens of male victims or those who identify as LBGTQ+. These groups tend to be underrepresented in trafficking research despite their known presence [182, 183]. Inclusivity of more diverse victims and trafficker demographics in trafficking research expands insights into the unique characteristics, needs and behaviors across trafficking.

Beyond the demographics of those impacted by trafficking, it is apparent that there is an opportunity to expand the diversity of the demographics of trafficking locations. Whereas trafficking occurs globally in various facets of society, and differs in its appearance across cultures, many of the sex trafficking studies we reviewed were conducted within developed countries (approximately 78.0%). Thus, a clear opportunity exists to conduct anti-HT research in developing countries.

It is altogether possible that the lack of research identified in these areas may be a direct result of constrained factors such as the limited amount of data available. Even so, greater research diversity in trafficking typology and demographics will allow for an improved understanding of the extent and impact of trafficking worldwide, as well as greater insights into how future research can help address associated needs in the fight against HT. This brings us to the next research gap observed: the need to diversity data sources.

Diversify the data sources. Data collection and analysis have helped in the fight against HT in many ways, including the identification of HT victims [72, 142, 149], informing prevention campaigns [150], and detecting trafficking network behaviors [54, 89]. While thorough data analysis lays the necessary groundwork for such discoveries, it relies upon the utilization of a variety of data from disparate sources.

Over 78.0% of all studies in the set used secondary data sources exclusively (Fig 1.6). The use of secondary data is common across many domains and proves beneficial given that the data already exists, is oftentimes publicly available, and can provide researchers with large amounts of data they might not be able to obtain otherwise. In the context of the OR and Analytics studies observed, over 22.0% of the secondary data sources used were from the same source, the now defunct, backpage.com [177]. Of the secondary data sources, around 54.0% of these studies focused on sex trafficking. So, while secondary data sources have their place, as noted in the previous section the lack of data diversity compounds issues such as the typologies studied. Obtaining more robust data can broaden the information available, provide better insights into the scope of trafficking (such as prevalence), and examine changes that occur over time. Even with the limited current state of available data, OR and Analytics can still offer valuable contributions.

Researchers, particularly those in the Analytics community, could leverage collaborations and focus on developing the already existing, practitioner-collected data and help to operationalize it. For example, a common issue faced in anti-HT analysis pertains to missing and incomplete data [184]. Many analytical methods exist that could help bridge this gap and improve the current anti-HT data landscape. Additionally, although many disparate datasets are available and growing, the anti-HT community rarely leverages combined data sources in a way to assess the status, trends, and dynamics of trafficking activities, and is another avenue for future work.

Despite the usefulness of secondary data, they have drawbacks that may hinder the results of a study. Secondary data is often collected with another objective in mind and therefore the nuances of that data collection process may cause bias. For example, what constitutes trafficking may differ from study to study and is a well-documented issue across anti-HT research [185–187]. There may also be a context in which data simply does not exist. For these and other reasons, we advocate for the collection and use of primary data where possible. While primary data may be more resource-intensive to obtain, it provides researchers with curated data for their specific research goals. When possible, sharing this data within the anti-HT community provides additional resources for other work. In addition, embracing collaborations with practitioners or other researchers in the anti-HT domain throughout the data collection phase can serve to increase the quality of this data by providing domain-specific insights. Building such collaborations demonstrates effective employment of the partnership principle (the fourth P) and embodies our next recommendation.

Better inclusion and collaboration of the 4Ps. The 4Ps (prevention, protection, prosecution, and partnership) are widely acknowledged as a holistic set of principles that accounts for the spectrum of anti-HT efforts. To date, the majority of OR and Analytics studies in the set appear to be focused on prosecution (Fig 1.6). Thus, while there exists demonstrated impact for prosecution-related activities, there are opportunities to contribute to anti-HT efforts in the spheres of prevention, protection, and partnership. A key way to increase the impact of OR and Analytics research in the fight against HT is to be keenly aware of all stakeholders involved, their various objectives, and how the research addresses the 4Ps. For example, while law enforcement may make decisions based on the likelihood of prosecuting traffickers, possibly at the expense of additional trauma to victims, Non-Government Organizations (NGOs) may focus more on the immediate needs of the survivors, offering an avenue for research around prevention and protection.

NGOs and governmental agencies often work directly with victims and survivors and could both inform avenues for profitable research studies, and themselves benefit from collaboration with OR and Analytics researchers. Given the often extreme resource constraints under which NGOs and governmental organizations operate, examining ways to evaluate current operations and improve resource allocation is a direction that deserves more study; less than 3.0% of all studies considered these areas.

Beyond the scope of the present study, OR and Analytics can help in the fight against HT by looking at push factors associated with HT. These areas include and are not limited to poverty, abuse, and lack of resources to meet basic needs [188]. More broadly, looking at ways to help improve circumstances of vulnerable populations at risk of HT is a needed avenue for future research.

While there are many ways in which the OR and Analytics communities can apply methods in the fight against HT, researchers ought to judiciously evaluate the problem context at hand, and whether an off-the-shelf method is justified; likely, the context warrants in-depth understanding, so that proper methodologies can be developed to accurately model and address the trafficking context [5].

Section 1.5

Conclusions

This survey provides a synopsis of the current state of the literature in OR and Analytics approaches in anti-HT contexts by surveying the research methodologies adopted in studies published from 2010 through 2021. A total of 142 studies were included in the set and examined, demonstrating the ability and promise of applying analytical methods to advance the fight against HT. A number of themes arose after careful review of the features of these studies, thereby illustrating opportunities for future research. We observed an increasing trend in the number of studies for both OR and Analytics, thus demonstrating a growing awareness of the issue of HT. However, the tendency of these works to focus specifically on sex trafficking underscores the need for future research in labor trafficking. Very few (less than 24.0%) of the studies on anti-HT in OR and Analytics focus on a specific sub-population, potentially failing to consider the diverse needs of victims and survivors. Existing OR and Analytics studies echo the anti-HT community at large for more available data. HT is diverse and nuanced, and researchers should make careful considerations when adapting existing methods to this vexing societal issue, considering efforts equally in prevention, protection, prosecution, and partnership.

Chapter 2

Estimating Effectiveness of Identifying Human Trafficking via Data Envelopment Analysis[1]

Section 2.1

Introduction

Human trafficking involves the commercial exchange and exploitation of humans for monetary gain or benefit and constitutes a gross violation of human rights [189]. In 2016, an estimated 24.9 million people globally were being trafficked for sex and/or labor [190]. Trafficking occurs both domestically and internationally in all regions of the world where individuals are made vulnerable by environments of poverty, conflict, natural disaster, unemployment, and desperation [191]. Victims range from child soldiers and child brides to domestic workers (e.g., housekeepers), forced laborers (in occupations including commercial fishing, manufacturing, construction, mining, and agricultural work, to name a few), people in the commercial sex industry, and beggars [192,193]. Forced labor and sexual exploitation generate an estimated $150 billion (U.S.) globally in illegal profits each year [194]. In addition to economic impacts, trafficking has severe health implications. Physical and sexual violence are common among survivors of sex trafficking, as is profound psychological manipulation. Labor trafficking victims often suffer from exhaustion, dehydration, heat strokes, hypothermia, respiratory issues, and skin infections [195]. Trafficked survivors may experience complex trauma as a result of forced isolation, deprivation, psychological abuse, degradation, and threats made to themselves or others [196].

[1]Dimas, G. L., El Khalkhali, M., Bender, A., Maass, K. L., Konrad, R., Blom, J. S., Zhu, J., Trapp, A. C. (2023). Estimating Effectiveness of Identifying Human Trafficking via Data Envelopment Analysis. Accepted, INFORMS Journal on Applied Analytics (Preprint available at https://arxiv.org/abs/2012.07746)

To address the pervasiveness of human trafficking, most countries have enacted laws and policies to prosecute those who traffic humans and to provide assistance and protection to survivors of human trafficking. Awareness of and efforts to combat human trafficking are increasing. Although anti–human trafficking efforts to date have largely focused on the sex trafficking of women, recent efforts have acknowledged that trafficking is a complex system that affects people of all genders and ages in a wide range of industries [197].

Despite the social, moral, and economic need to address human trafficking, resources to address this issue remain extremely limited [198]. Although the gap between the current funding of antitrafficking interventions and needs is challenging to quantify, the literature consistently refers to the inadequacy of current funding for human trafficking interventions [199–201]. A critical challenge for agencies and governments is to use scarce resources efficiently, as well as to effectively evaluate the impact of their antitrafficking initiatives. Even so, many anti–human trafficking interventions continue to operate without an adequate evidence base [202, 203]. Although the antitrafficking community largely lacks the analytical expertise and resources necessary to evaluate the effective allocation of resources, there is a tremendous opportunity to (1) increase quality evaluations of anti–human trafficking programs; (2) ensure that programs are targeted, implemented, and delivered effectively; and (3) improve the knowledge concerning the impact of interventions [203].

Data analytics can be used to improve the lives of individuals [204, 205], support decision making in practice [206], and prove the value of routine data collection for operational decision making [207, 208]. Despite evidence suggesting that the awareness of the value of routinely collected data may be increasing [209, 210], the benefits gained from data analysis in the nonprofit and humanitarian sectors are rarely documented. More alarmingly, strategies for data collection and management to support such analyses are borrowed from the for-profit sector and do not address the unique needs of humanitarian nonprofits, such as collecting data regarding vulnerable populations in complex environments with limited funding. Furthermore, data in nonprofit and humanitarian sectors are not conducive to clean or simple data collection. In the severely resource-constrained environments in which nonprofits typically operate, data analyses to support operational strategy and decision making can have positive effects such as improving finances, increasing operational efficiency, and enhancing donor relationships. Although many donors "require an assessment of the deployment and performance improvements resulting from their investments" and continuous improvement, it can be difficult to develop the processes, structures, and systems necessary to support strategy and decision making because nonprofits typically do not have the means [211]. Routinely collected and analyzed data present a valuable and underutilized opportunity for nonprofits.

Performance assessment and monitoring is critical to establishing benchmarks for best practices and guiding improvement recommendations. Data envelopment analysis (DEA) is a data-driven analytical tool used to assess the performance of units within an organization, or across organizations. Such analytics can highlight possible improvements in effectiveness for organizations, including those involved in antitrafficking operations. We use DEA to analyze existing data and evaluate the performance of border stations of nongovernmental organization (NGO) Love Justice International (LJI) in Nepal engaged in the trafficking intervention strategy known as transit monitoring. The performance of these stations is evaluated for their effectiveness at intercepting potential human-trafficking victims given the amount of resources (staff, etc.) available. These data enable comparison of border stations to determine which are efficient relative to other decision-making units (DMUs). Understanding the relative efficiency of border stations enables recommendations for best practices for improving operations. Operations research methods applied in the context of anti–human trafficking are scarce, and this study, to the best of our knowledge, is the first to use DEA [6]. In so doing, we illustrate how analytical methods can be applied to operational decision making in the nonprofit environment.

In what follows, we overview how transit monitoring is used as a strategy to combat human trafficking and other public sector contexts in which DEA has successfully been applied. We then describe the data and DEA approach, homogeneity criteria, model inputs and outputs, and model results. We conclude with a list of actionable recommendations and discuss the limitations and generalizability of the model. Additional model and data details can be found in the appendices.

Section 2.2 Transit Monitoring

[212] depict trafficking as a series of event stages during which risks to an individual and intervention opportunities may arise: recruitment, travel-transit, exploitation, detention, integration, and retrafficking. Although human trafficking, by definition, does not require movement of victims from one location to another, travel occurs in many human trafficking contexts and provides an opportunity for intervention.

The objective of transit monitoring is to identify and intercept those at risk of being trafficked in the travel-transit phase and before the exploitation phase [212, 213] so as to limit the level of trauma that any victims might experience. Trained personnel located along trafficking routes (such as transportation hubs and state border crossings) assess trafficking indicators and engage the potential victims in transit, involving government authorities when appropriate. Transit monitoring is more nuanced than the broader concept of network interdiction. Transit monitoring focuses on interrupting trafficking *before* exploitation, whereas interdiction does not make this distinction. Although it is possible to optimize strategic resource-allocation decisions for network interdiction (including trafficking networks), the focus of this study is on the use of DEA to evaluate the effectiveness of LJI staff at conducting transit-monitoring activities.

2.2.1. Human Trafficking and Transit Monitoring in Nepal

Human trafficking in Nepal is a growing criminal industry [191]. Nepal is considered a "source country" of men, women, and children subjected to labor and sex trafficking [191]. The interaction of poverty, development, and relevant policies affects the vulnerability of its population to trafficking [191, 214]. A large proportion of the population is estimated to be unemployed (42%) or living below the poverty line (38%). As of 2014, the most recent figures available, Nepal ranks 142nd out of 189 countries on the Human Development Index [215]; it has a low literacy rate (66%), and more than 80% of the country's inhabitants live in rural areas [215]. Most trafficking victims in Nepal are female adults and children who not only suffer from gender inequality but are also particularly vulnerable because of economic insecurity, recent national disasters, and poverty [216–218]. Such factors generate high levels of migration to urban centers in Nepal, India, and the Middle East, where trafficking victims are lured with promises of a better life, jobs, and false marriage proposals or through the coercion of indebted families to sell their children [218–220]. Migration is generally defined as the voluntary movement of persons within or across borders in search of a better livelihood. Trafficking and exploitation are associated with migration in two aspects: a person may willingly migrate for employment but find the work conditions to be exploitative or may be deceived regarding the kind of work they will be doing [216, 221].

In 1996, an antitrafficking operation on Indian brothels identified 200 female Nepali minors being trafficked. This occurrence spurred increased attention of anti–human trafficking in Nepal. The Nepali government's failure to help victims recover and rehabilitate motivated seven NGOs to take action [213]. Since 1996, these NGOs have played a significant role in combating human trafficking in Nepal, most notably by establishing transit-monitoring and/or border-monitoring stations along the largely unregulated Indian border. Although trade between Nepal and India is monitored through 22 checkpoint stations, Indian and Nepali individuals are permitted to cross the border at any point. (However, Nepal temporarily closed its international borders in 2020 because of the COVID-19 pandemic [222].) In comparison, citizens from other countries are only permitted to cross at six border checkpoint stations after obtaining entrance and exit visas [223]. Unfortunately, the negligence and corruption of the border-monitoring protocol have led to a prominent culture of transnational crime that includes human trafficking [224]. In 2018, an estimated 171,000 Nepali individuals were victims of human trafficking or forced marriage [225]. The accuracy of the number of individuals trafficked in and out of Nepal is disputed, and the lack of precise and accurate data is detrimental to measuring Nepal's progress in combating human trafficking [213].

LJI, formally known as Tiny Hands, is an NGO that combats human trafficking through transit-monitoring methods. LJI focuses on preventing human trafficking before exploitation by placing trained personnel at key transit points to identify people who exhibit risk indicators associated with currently being trafficked or being trafficked in the near future. Although LJI operates in 15 countries, the DEA model we present was developed for LJI operations in Nepal where transit-monitoring practices were first launched in 2006 along the Nepal-India border. Over the past 14 years, LJI has grown to an operation of nearly 30 monitoring station locations throughout Nepal. Most stations are located on the Indian border, but some are located in interior major transportation hubs, such as regional bus stations, and airports.

The DEA model enables LJI to identify both high- and low-performing stations, where performance refers to a station's effectiveness of intercepting potential human-trafficking victims given the amount of available resources (such as staff). LJI is using the results of this study to develop best practices, evaluate and improve performance at inefficient stations, and, more generally, use their limited antitrafficking resources more efficiently.

Section 2.3 DEA Applications in the Public Sector

Few applications of analytical approaches for anti–human trafficking operations currently exist (see, e.g., [6, 226–228]). To the best of our knowledge, there are no known DEA applications in antitrafficking operations. DEA is useful to evaluate the relative efficiency of units within an organization that has different levels of inputs (e.g., resources, demand) that are used to produce outputs (e.g., productivity measures). *Efficiency* refers to a unit's ability to produce an expected amount of output given the amount of input resources. DEA models are employed in areas that range from the assessment of public organizations (such as healthcare systems, educational institutions, and governmental bodies) to private organizations (such as banks, restaurants, and service providers). [229] provide a recent review of the development and use of DEA models in the public sector. DEA has been applied to humanitarian efforts, such as the work done by [230], who measured the effectiveness of humanitarian aid efforts across 106 different countries. Similarly, [231] consider the efficiency of humanitarian supply chains using a network DEA model.

In a domain loosely related to transit monitoring, DEA has been used in policing to evaluate operational efficiencies. Aspects of policing work have some parallels with transit monitoring; namely, units of an organization or agency are seeking to identify illicit behavior. For example, [232] evaluated the relative efficiency of Taiwanese police precincts and found that operational and process differences primarily depend on the resident population and location. [233] measured the efficiency of state police units in India and found that a DEA model can generate targets of performance, identify inefficient departments, and determine adequate levels of operation and improvements in the units of criminal justice systems. Through DEA, [234] assessed the effectiveness of the Spanish police force to determine the most effective overall unit. Using an adjusted version of the DEA methodology "Benefit-of-the-Doubt," [235] advocate for the use of a custom-made operations research framework to evaluate citizen satisfaction with police effectiveness of community-oriented local police forces in Belgium.

DEA has also been used to measure the performance of nonprofits [236, 237]. DEA emerged as an alternative method to measure the operational efficiency of nonprofits because it is capable of handling multiple performance metrics that are categorized as "inputs" and "outputs" to provide a single composite efficiency score for each nonprofit organization compared with other organizations that produce similar outputs. [238] evaluated the efficiency of five NGOs in Turkey dedicated to serving people suffering from wars, invasions, and natural disasters and supporting displaced refugees while distributing other forms of aid without discrimination. [239] developed a DEA model to measure the efficiency of microfinance institutions and alleviate global poverty by extending financing to the poor for social and financial efficiency. In the second stage of this study, [239] compare the performance and factors that contribute to the efficiency of Islamic microfinance institutions. A challenge arising from evaluating efficiency for nonprofits is that NGOs have a social objective, meaning they do not take part in competitive markets that use net income and rates of return as efficiency indicators [240]. Therefore, many have used arbitrary measures of efficiency that involve developing weighting methods to quantify the outputs of a nonprofit [240].

Section 2.4

Data

LJI staff who conduct transit monitoring are trained in a multistep process for potential victim identification: (1) visual identification of people traveling near transit stations to detect suspicious activity; (2) engagement with suspicious parties for heightened profiling to obtain more specific details, should this rise to the level of suspected trafficking (either currently or in the near future); (3) interception and further questioning of potential victims or traffickers, culminating in a possible completion of an Interception Report Form (IRF); (4) completion of a Victim Interview Form (VIF) for each individual in the party; and (5) verification of the responses to help validate pertinent details. The verification processes usually require cultural knowledge, cross-checking data in LJI's "Fusion Center" database (consisting of 9,000 individuals known or suspected of being involved in human trafficking), and following up with a third party by phone, such as parents, relatives, universities, or employers [213].

After visual identification of suspicious activity, trained LJI staff intercept the individual(s) and administer an IRF. The IRF consists of a point-based system, called "red flags," created by LJI to help guide staff in determining an occurrence of suspected trafficking. This point system assigns larger numeric values to questions more indicative of trafficking risk. For a more thorough explanation of this process, we refer the interested reader to [213].

When a party's responses to questions on the IRF exceed the predetermined "red flag" threshold, trained LJI staff administer a VIF for each individual in the group [213]. LJI staff may then refer the potential victim and perpetrator, if present, to local law enforcement for follow-up investigation and prosecution. If in LJI's judgment, the "red flag" threshold is not exceeded and there is insufficient support for an occurrence of suspected trafficking, a VIF is not filled out, and the party continues their travels without further intervention. LJI replaced the VIF with a similar form called the Case Information Form (CIF) in July 2018; for ease of reading, we hereafter jointly refer to these forms as VIF.

If *both* an IRF and a VIF have been completed for an individual, these paper forms are translated from Nepali to English and sent to data analysis staff at LJI. The data entry specialists enter IRF and VIF form data into a standard Microsoft Excel workbook with rows representing observations (potential victims) and columns representing various features of interest (responses from the respective forms).

We had access to over 8,700 IRFs and nearly 4,500 VIFs ranging from late 2011 to 2021. Collectively, the IRF and VIF data consist of over 350 data fields and capture both potential labor and sex trafficking transit activity. The data include detailed demographic information on the victims (age, gender, caste, health, marital status, history of abuse) and on the traffickers (age, gender, occupation, and how they met the victims). Rich information is also available regarding the trafficking supply networks, including victim origin and destination, recruitment methods, transit duration, methods of transportation, and safe houses. For example, Figure 2.1 depicts the originating districts of potential trafficking victims who were intercepted at LJI border stations in 2017 (for ease of reading, only the 10 LJI border stations with the highest recorded transit activity are displayed).

Before data analysis and development of the DEA model, the data needed to be cleaned. This process involved linking the IRFs with their corresponding VIFs, organizing the data in a manner to reduce any redundancies, filling in and removing missing or incomplete data, and completing additional general data cleaning steps. Next, the data were filtered to find stations matching our homogeneity criteria, and then each of the input and output features was aggregated on the quarter level as described in the methodology section (see Table 2.1). This process was conducted in the R scripting language.

Section 2.5

Methodology: Data Envelopment Analysis

We employed DEA to quantify the relative efficiencies of deploying transit monitoring at various LJI stations in Nepal. The purpose was to identify which LJI stations in Nepal are considered efficient and how the level of

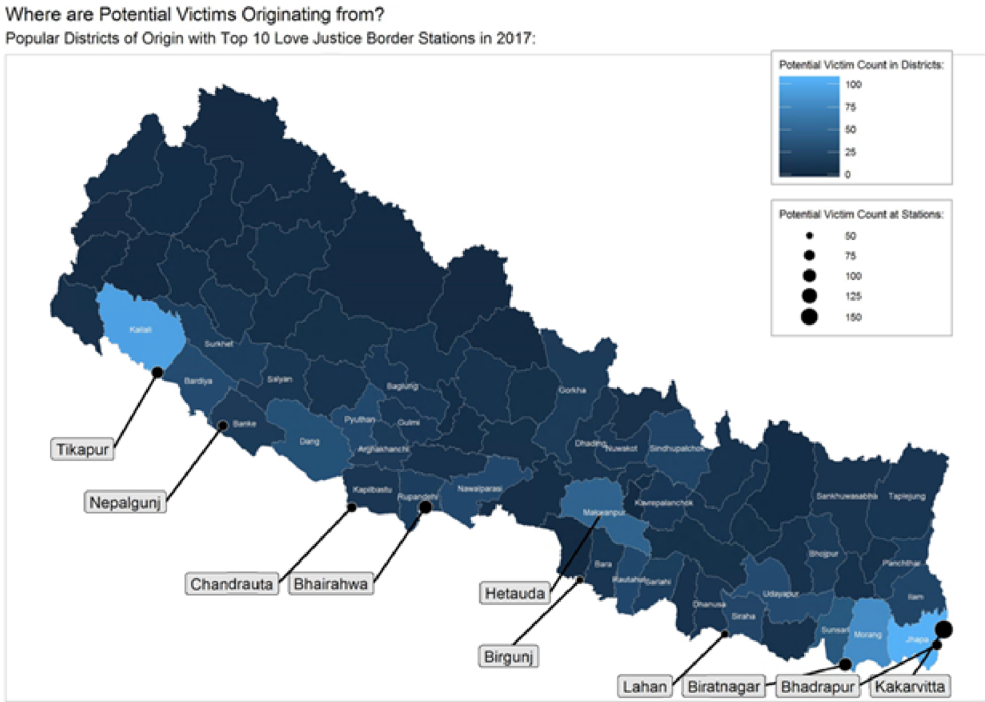

Figure 2.1: **The Districts of Origin and the Top 10 LJI Border Stations, by Frequency of Victims, for 2017.**

efficiency ranks among the stations. These results could then be used to identify potential methods of improving human trafficking transit-monitoring tactics.

DEA is a mathematical programming method used to compute the relative efficiency of multiple DMUs, such as a hospital in a healthcare system, which have different inputs and outputs [241]. The performance of each DMU is evaluated by assessing its ability to produce associated outputs by consuming its inputs, with respect to all other considered DMUs [242].

DEA models are commonly categorized by the shape of the identified best-practice frontiers. In the DEA literature, we use variable returns to scale (VRS) and constant returns to scale (CRS) to classify two basic DEA models. In the context of LJI transit monitoring, because the number of inputs does not result in a proportional change in outputs, a VRS model was most appropriate (see Appendix A.1 for details). Moreover, our objective is to maximize the outputs achieved, and thus, we employed an output-oriented VRS model [243]; Appendix A.1 details the algebraic formulation and further motivates our use of the VRS model. We additionally conducted a cross-efficiency analysis that allows for a quantified ranking of the *relative* efficiencies of DMUs, which extends the findings of a standard DEA investigation that only identifies efficient or inefficient stations in a binary fashion [242].

It is generally required that all DMUs chosen must be homogenous under DEA. After all, as DEA aims to view all stations in their best relative light, it is only fair to compare DMUs that operate under relatively similar environments. In our context, the criteria we consider for homogeneity are (1) the presence of a station manager,

(2) the quality of data, and (3) flow through station locations. Stations that shared similar ranges for these parameters were grouped. Each station considered for evaluation was required to have a station manager because these stations tend to be better managed and share similar operational standards. The stations chosen were also required to have data (positive counts of IRF and VIF forms) for all consecutive quarters during our timeline to ensure reliability of the data. Finally, we grouped stations with the same flow together for consideration. The station flow is the magnitude of total travelers through a station estimated from several factors, including the distance from the Indian border, number of airports, number of bus or train stations, number of major highways, and population density. Based on their scores in each category, a value of 1, 2, or 3 was assigned to the stations, with 1 and 3 representing the lowest and highest estimated flows, respectively. For this analysis, we considered only stations with a flow of 3, under the assumption that larger flows of people would lead to an increased likelihood of intercepting trafficking activity. We validated our estimated flow ranking with our partners at LJI, which was well received. Out of the 18 stations for which we had data, 7 met these standards. Table 2.1 summarizes the selection process for these stations. The seven selected stations for our analysis are Mahendranagar, Nepalgunj, Birgunj, Kakarvitta, Biratnagar, Bhairawa, and Bhadrapur. Figure 2.2 displays these stations geographically.

Table 2.1: **The DMU Selection Criteria and Process, Resulting in Seven Selected Stations for Further Analysis.**

Step	Selection criteria	Total stations meeting criteria
0	All stations operating between 2011 and 2021 Number of stations observed in our data set	18
1	Presence of a station manager Number of stations with a station manager on staff	12
2	Quality of data Selected stations that operated consistently over time (that is, have positive IRF and VIF form counts)	11
3	Flow through station location Stations with a high level (3) of flow	7
Final data set: 7 stations * 13 quarters = 91 total DMUs		

To ensure a sufficient number of DMUs [244], each of the seven stations was expanded to 13 calendar quarters of data over which we had complete information (from Q2 2016 through Q2 2019). This created 91 station-quarter DMUs, satisfying the empirical rule that suggests the number of DMUs be at least twice the number of inputs and outputs combined [245]. Because we use DEA as a tool for relative performance evaluation, rather than a production function estimate, the number of DMUs is much less critical for our study (see, e.g., [245]). For ease of reading, we refer to “station-quarters” as “stations” where unambiguous. Although our original data set contained information from late 2011 to 2021, consistent, high-quality data for these seven stations were only available for four years; the 1st quarter in our analysis represents April to June of 2016, and the 13th and last quarter in our

Figure 2.2: **The Seven LJI Border Crossing Stations Meeting the Qualifications for Inclusion in Our Analysis.**

analysis represents April to June of 2019. We note that the COVID-19 pandemic impacted operations during 2020 and 2021, and therefore, the data were incomplete and excluded from our analysis.

DEA was applied to the 91 LJI interception station-quarters to evaluate their performance in applying transit monitoring. By assessing the efficiency of the station-quarters using the inputs consumed and outputs produced, LJI can begin to understand which station-quarters performed better at following procedures and intercepting individuals at risk for human trafficking. Such insights can then empower further root-cause analyses, which in turn provide LJI with a blueprint for improving overall station performance.

2.5.1. DEA Formulation to Evaluate LJI Transit-Monitoring Stations

Through discussions with LJI and availability of related data, we identified the key inputs and outputs for our DEA modeling to evaluate the performance of human trafficking interception stations. We provide details of our final model and refer the reader to Appendix A.2 for an analysis of other considered inputs and outputs.

Inputs. We consider three inputs for each station: *number of staff*, *test scores*, and *hours worked by staff*. The number of staff members represents the average number of LJI personnel employed by each station within a quarter, which ranged from 2 to 12 individuals. Test scores represent the average staff performance on border knowledge examinations, a mandatory test each staff member must take before employment. The average quarterly test scores of the staff assigned to a station ranged from 20 to 25. The average weekly hours worked by staff ranged from 36 to 50.

Outputs. We consider four outputs for each station: *number of suspected trafficking occurrences* (count of IRFs), *number of potential victims* (recorded on VIFs), *IRF completeness*, and *VIF completeness*. There is one VIF for each individual recorded on the corresponding IRF form that exceeded the predetermined trafficking risk threshold. The completeness of the IRFs and VIFs represents the thoroughness of staff filling out all of the minimum details necessary in each respective form.

In consultation with LJI staff, station IRF and VIF completeness were calculated as the average percentage of required questions filled out in each form. If these questions were answered, a value of 1 was assigned to that question for that form; otherwise, a value of 0 was assigned. The completeness of each form is then calculated by taking the percentage of the required questions *answered* on a form out of the total questions required. Because the IRF asks questions regarding the collective group of people crossing the border together, a single IRF may contain information for multiple people. However, the VIF is completed individually for each person in the group. Because a group of people crossing the border generates a single IRF and multiple VIFs, the VIF completeness measure associated with the encounter can be interpreted in multiple ways. For example, if three people are traveling together in a group, the VIF completeness for the encounter could be measured by the minimum, maximum, average, or sum of the completeness of the VIF forms for the three individual people. In consultations with LJI stakeholders, it was determined that an encounter's VIF completeness was most appropriately defined by averaging the VIF completeness of all individuals in the group.

Section 2.6

Findings and Recommendations

In LJI's context, efficient stations are ones that have an adequate number of IRFs and VIFs filled out with a sufficient level of completeness for each quarter given (1) the average number of staff employed at the station, (2) the staff's border knowledge test scores, and (3) the average weekly hours worked by staff. A station is considered inefficient if it fails to produce adequate output from its characteristic input, indicated by a DEA model score of less than one. Figure 2.3 illustrates the results of our initial DEA investigations on station efficiencies. Efficient stations are indicated by yellow circles, whereas blue diamonds indicate inefficient stations. Although none of the seven stations were efficient every quarter, Kakarvitta and Biratnagar were efficient for a majority of quarters, with Kakarvitta having only 2 out of 13 quarters identified as inefficient. The remaining five stations were inefficient for over half of the quarters.

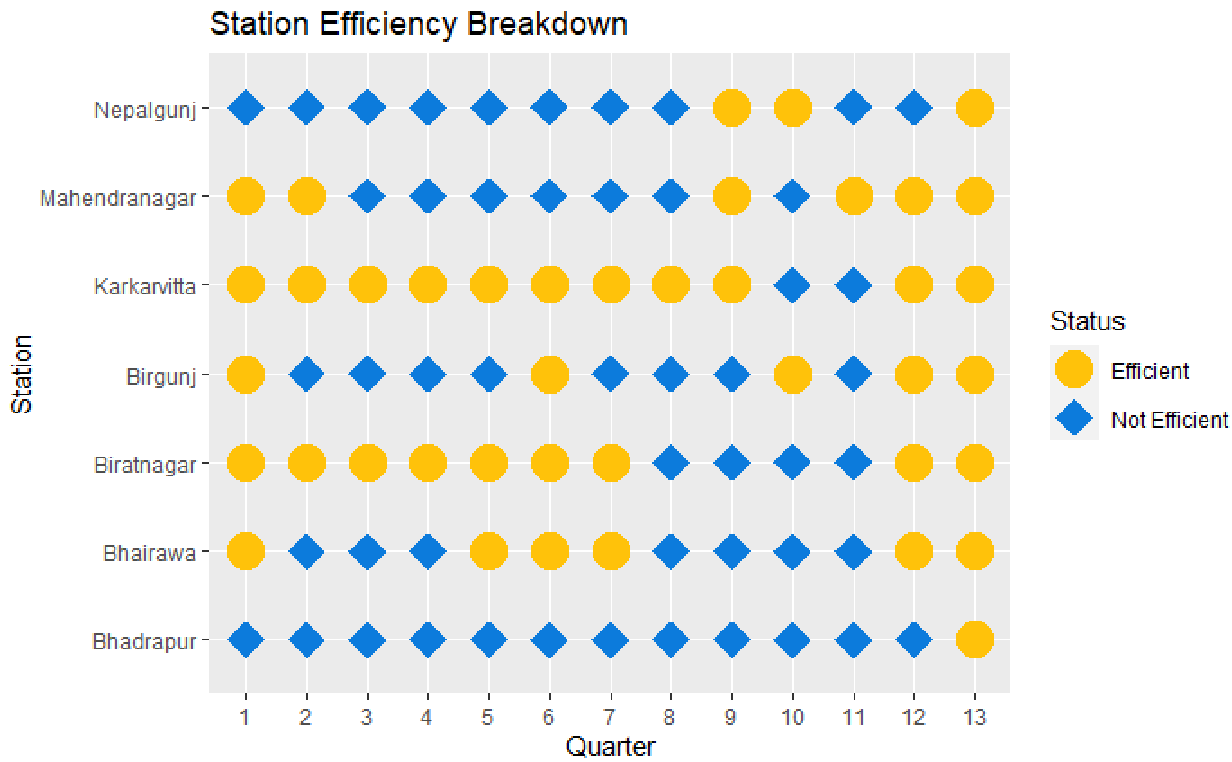

Figure 2.3: **Efficiency, by Station and Quarter, with Respect to the Best-Practice Frontier.**

Of interest is that all of the stations were efficient in Quarter 13 (Q2, 2019). Although at first glance one might conclude that all stations have reached a long-term efficient state, this is an incomplete assessment; variability across quarters reveals that efficiency can readily change from quarter to quarter.

Figure 2.3 reveals DMUs on the best-practice frontier over 13 quarters and provides initial insights into a station's own efficiency trajectory. Because station efficiency may vary over quarters, it would be incomplete to infer from Figure 2.3 the best-performing stations. To understand the performance of individual stations with respect to others, we calculated cross-efficiency values under VRS [242]. Cross-efficiency is a method to rank DMUs using weights that considers both peer and self-evaluation and results in a ranking of stations based on their average cross-efficiency values over time, as compared with every other station. The average cross-efficiency–based ranking for all seven stations is shown in Figure 2.4. We note that although every station had at least one quarter of inefficiency, the average cross-efficiency for each station is at least 0.856.

Taken together, Figures 2.3 and 2.4 offer complementary views on station efficiency. Figure 2.3 reveals quarter-by-quarter best-practice frontiers across the seven stations, whereas Figure 2.4 reveals cross-efficiency scores for each station averaged across all quarters. The resulting differences in Figures 2.3 and 2.4 with respect to rank and best-practice frontiers underscore the differences that can result from two different, yet related, comparison methodologies.

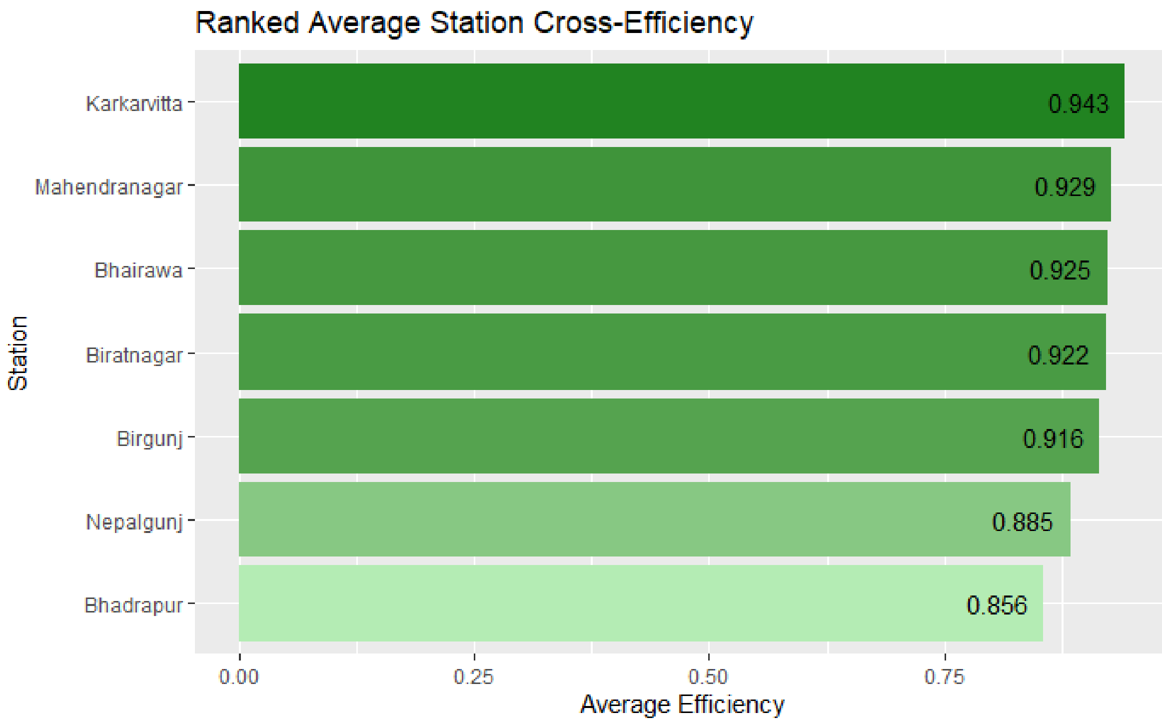

Figure 2.4: **Cross-Efficiency, by Station, Averaged over Quarters.**

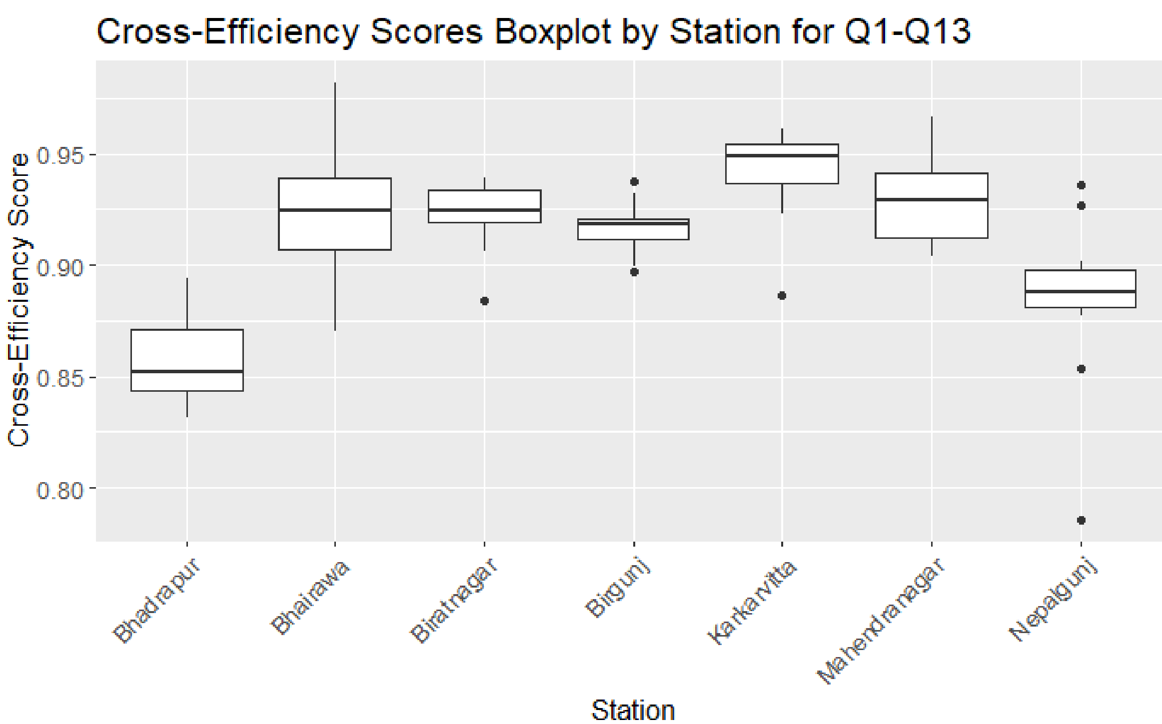

Figure 2.5: **Cross-Efficiency by Station, with Boxplots Depicting Distribution over 13 Quarters.**

Figure 2.5 shows quarter-to-quarter variation in cross-efficiency scores. Biratnagar, Birgunj, Karkarvitta, and Nepalgunj have the smallest interquartile ranges, demonstrating low variance and consistency in performance. Bhadrapur, Bhairawa, and Mahendranagar have larger interquartile ranges, indicating more performance variability between quarters. Figure 2.5 also substantiates that Nepalgunj and Bhadrapur consistently produced the lowest cross-efficiency scores.

Through careful interpretation, the results of a station's DEA performance can inform implementation toward the best-practice frontier, which can be seen as the threshold in which a DMU is considered efficient based on its inputs and outputs. A station along the best-practice frontier indicates it is deemed efficient because of its ability to produce outputs with the given inputs. Based on our findings, we developed recommendations to increase the consistency of a station's performance efficiency. These findings are based on numerical analysis from the DEA outputs, observations, and discussions with LJI. The numerical analysis refers to the DEA outputs that describe the percentages that other efficient stations are contributing to the efficiency score of inefficient stations. The efficient output values are the dot product of the efficient stations' output values and their percentage contribution of inefficiency. The DEA model results also disclosed the number of additional units of outputs that each station needs to produce to be considered efficient.

Figures 2.6 through 2.9 show the results of this analysis for each output feature. The size of the circle depicts the percent of additional units required for a DMU to have been considered efficient. The absence of a circle indicates the outputs of a station were already operating on the best-practice frontier. For example, Figure 2.6 shows that Nepalgunj requires an additional 2.5% of IRF forms to be completed in Quarter 11 to reach the efficiency frontier. These plots quickly reveal which stations require the largest increase in outputs. Based on these values, we summarized observations from the numerical analysis for each station in Table 2.2. We provide our recommendations by highlighting (1) which stations require the most attention overall as well as (2) which outputs from each station need the most attention. Because of the limited amount of resources typically available for NGOs such as LJI, these recommendations help provide insights on where to prioritize their resources and in which areas (outputs). For each station (across all 13 quarters), we ranked which required the largest increase in outputs (Table 2.2, column 2). In addition, for each of these stations, we ranked which outputs require the most attention (Table 2.2, columns 4 through 7) and summarized these results (Table 2.2, column 3). The ranking of the outputs ranges from 1 to 4; some outputs have equal rank and therefore may have the same value. For example, looking at the station Nepalgunj (Table 2.2, row 1), we see this station requires the most improvement across all outputs. The output with the largest potential for improvement was the number of VIF forms, followed by both VIF completion and IRF forms, whereas IRF completion was the least important.

To improve the performance of stations that lie outside of the best-practice frontier because of VIF and IRF completeness (Table 2.2, columns 5 and 7), LJI should consider conducting additional IRF and VIF training and quality control checks on these stations. Recently, LJI has started to provide workers with a "cheat sheet" that

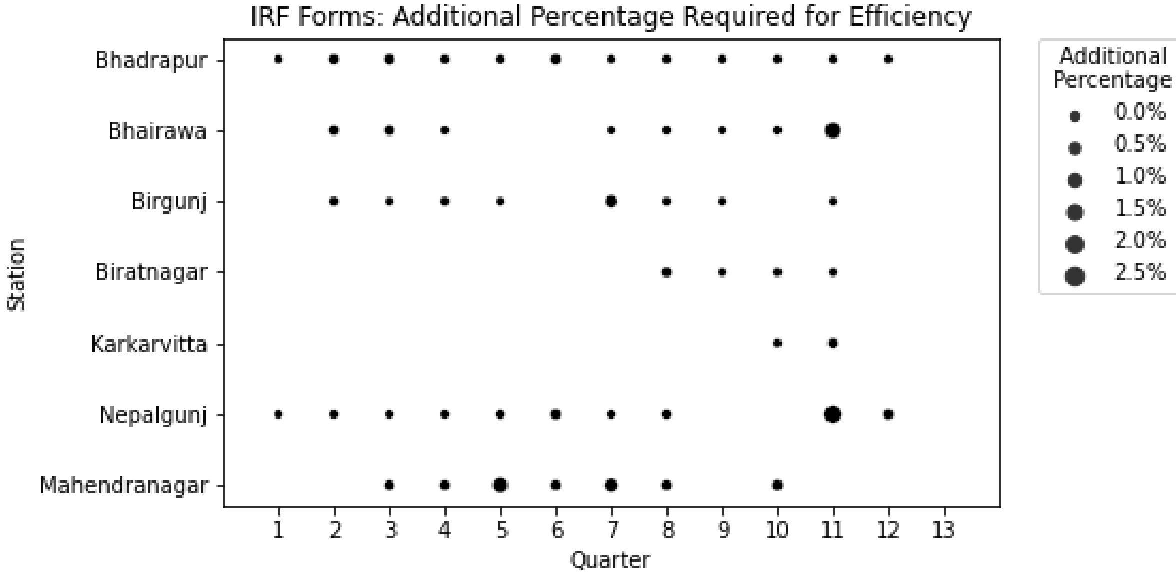

Figure 2.6: **The Additional Percentage of IRF Forms Needed for Efficiency Varies by Quarter and Station.**

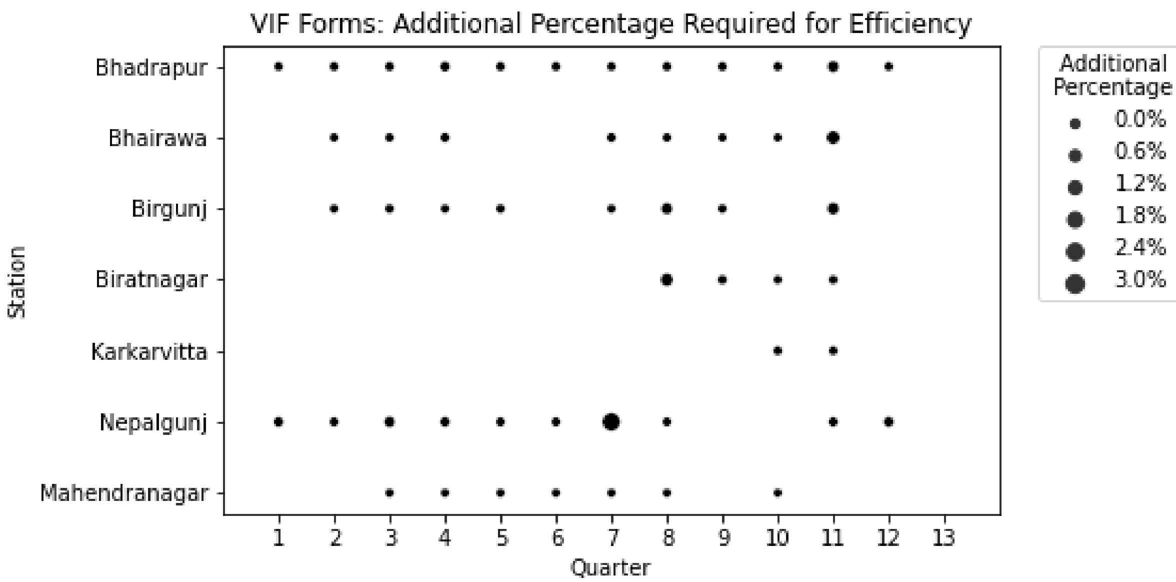

Figure 2.7: **The Additional Percentage of VIF Forms Needed for Efficiency Varies by Quarter and Station.**

emphasizes required questions on the IRF and VIF forms. The implementation of this new tool is expected to be an effective way to improve these outputs.

If the rate of trafficking were directly proportional to the rate of flow at all border stations, then *mathematically* speaking, increasing IRF and VIF collection suggests that LJI would intercept more individuals at risk of being trafficked—that is, fewer potential victims would be missed (Table 2.2, columns 4 and 6). As is so often the case in anti–human trafficking work, knowledge of the ground truth of actual trafficking cases is lacking. Hence, we instead used a proxy of overall flow, which represents not only trafficking occurrences but also legal transit—

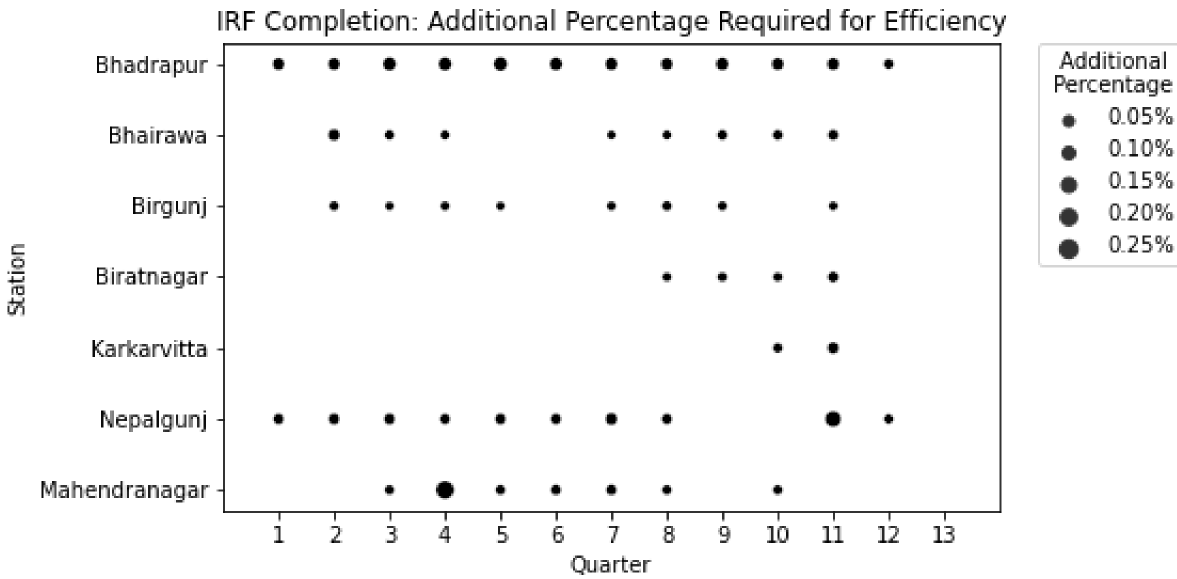

Figure 2.8: **The Additional Percentage of IRF Form Completion Needed for Efficiency Varies by Quarter and Station.**

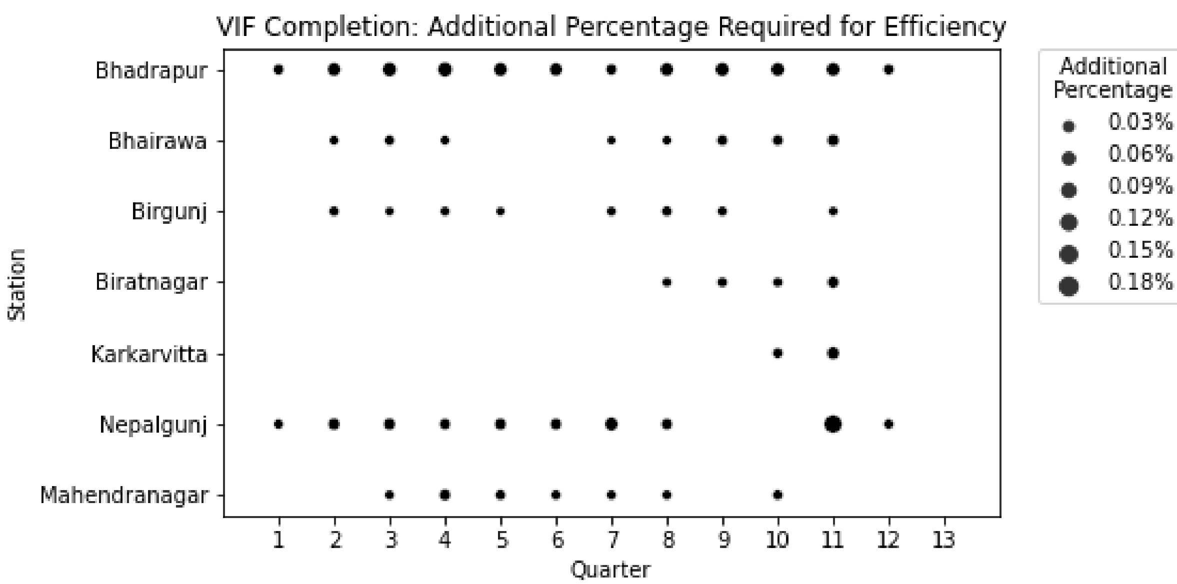

Figure 2.9: **The Additional Percentage of VIF Form Completion Needed for Efficiency Varies by Quarter and Station.**

ordinary transnational travel, including those willingly traveling for gainful employment.

As such, care is required in interpreting recommendations before putting them to practice. Rather than simply increasing IRF and VIF collection on the entire flow of travelers, it should be done *conditional on actual occurrences of trafficking.* We recommend that LJI conduct supplementary analysis to determine whether any of these stations are expected to have a less-than-proportional number of potential trafficking victims passing through their station relative to the total number of travelers. Stations with a less-than proportional trafficking flow may not actually

Table 2.2: **Key Areas for Improvement, by Station, Yielding Recommendations for Inefficient Stations Toward Best-Practice Frontier.**

		Recommendations				
Station	Overall priority	Key areas for improvement	VIF count	VIF completeness	IRF count	IRF completeness
Nepalgunj	1	VIF forms	**1**	2	2	3
Bhadrapur	2	VIF and IRF completion	2	**1**	2	**1**
Mahendranagar	3	IRF forms	4	3	**1**	2
Karkarvitta	4	VIF completion	3	**1**	2	2
Bhairawa	5	IRF forms	2	4	**1**	3
Biratnagar	6	VIF forms	**1**	2	4	3
Birgunj	7	VIF forms	**1**	2	4	3

need to increase the number of IRF and VIF forms collected because the low performance on this measure may be an artifact of fewer people being trafficked through this location.

That said, for stations where it is reasonable to assume that identification of additional potential trafficking victims will move them toward the best-practice frontier, we recommend LJI increase IRF and VIF collection on this targeted population of flow—actual trafficking occurrences—by improving strategies and tools for identification as well as continuing education and training for staff. Specifically, LJI should focus on whether there are any cultural or geographical factors, such as recent changes in migration trends; a newly favored route or terrain to cross the border in nearby locations that LJI staff have been unable to monitor; or ways traffickers may be coaching victims on responses to LJI questions to evade detection that may be affecting the number of potential victims identified at stations. Although such additional analyses are outside the scope of the present study, they can provide useful supplementary information to LJI regarding where to focus their efforts for further analysis.

Section 2.7

Conclusion

Human trafficking is a complex societal issue requiring a variety of approaches to help counter the social, health, and economic impacts associated with this crime. Transit monitoring has emerged as a promising strategy in reducing potential trafficking and exploitation activities, in which trained personnel are strategically located along potential trafficking routes to dynamically assess trafficking indicators and intercept probable victims before exploitation. Although transit monitoring is an effective strategy in preventing trafficking, transit-monitoring stations can vary in their performance, and there are opportunities to evaluate and improve stations by sharing best-practice recommendations among stations. To this end, we used DEA to differentiate between efficient and inefficient stations over time. Our approach is innovative as a performance management methodology for nonprofits in the antitrafficking sector because it allows decision makers to evaluate their organizations and units with multiple inputs and outputs, benchmark units against comparable peers, and plan for the future based on a realistic allocation of resources. It is believed that this effort is the first application of DEA to evaluate performance in anti–human trafficking efforts.

In resource-constrained environments in which nonprofits often operate, analyses to support operational decisions offer an opportunity to improve operations. Although such analyses have a long history of leading to more efficient operations in the for-profit sector, they appear less frequently in the nonprofit sector. This study presents the experiences from a collaboration between operations research analysts and a nonprofit engaged in anti–human trafficking initiatives. We discuss the use of DEA to examine the performance of stations engaging in transit monitoring in collaboration with Love Justice International, a nonprofit human trafficking organization engaged in transit monitoring along the Nepal-India border. The approach and results were evaluated in several meetings with LJI stakeholders, which provided insights for better calibrating our model with reality. This iterative process established our final model and highlights the significance of engaging all parties in determining model parameters (such as inputs and outputs) to increase the relevance of model outcomes (such as performance evaluation). Our analysis identified performance inefficiencies of individual transit-monitoring stations. A repository of all R and Python code used to create the DEA model and data visualizations can be found publicly at `https://github.com/gldimas/Nepal.HT.DEA`.

This study is not without limitations, which may also present as opportunities for future investigations. We acknowledge that variables that were not included in our modeling may affect its effectiveness. For example, other transit-monitoring NGOs operate at certain transit-monitoring stations, which could impact the inputs and outputs in our DEA model. We attempted to mitigate any confounding variables by basing our analysis on LJI's consistent data format. Although there may be other factors to consider, we were able to develop a framework that can be expanded on to help NGOs antitrafficking efforts.

Nonprofits often have an abundance of routinely collected data. However, such data often lie idle in locations that are inconvenient to access and formats that are challenging to manipulate. Moreover, data quality issues may exist. For these reasons, there is a great opportunity for future research on improving the collection and management of data in the nonprofit and NGO sector. By doing so, NGO and nonprofit sector data can be better leveraged, providing an immense opportunity for improved operational analyses that are both accurate and meaningful.

Although DEA produces meaningful recommendations, its mechanism and logic may prove challenging to grasp for stakeholders lacking an analytical background. We used visual aids and verbal explanations over several NGO stakeholder meetings to illustrate our findings. DEA does have methodological drawbacks, such as difficulty in handling negative numbers, challenges with handling zeros and missing data in a straightforward manner, and high sensitivity to outliers and bad data [246]. Some of these challenges were present in our study; as a result, some DMUs lacking positive counts of IRF and VIF forms during our timeframe needed to be omitted, and only a subset of stations could be evaluated. Additionally, our decision to add quarters to increase the number of comparison DMUs made our analysis unique in enabling quarter-by-quarter comparisons of stations with themselves. Although the use of quarter-by-quarter comparison increased the number of DMUs for analysis, the lower number of stations restricted the types of DEA analysis used in this work and is a limitation. Another limitation of our analysis is that the most recent data we used were over a year old because of impacts of the COVID-19 pandemic. Although this lag implies that more timely findings and actions are possible, the intent of the study was accomplished: to build a framework on which a nonprofit could measure performance and identify areas for improvement. The results of the DEA model confirmed anecdotal experiences of LJI operations. LJI is now empowered to take these results and this approach to identify the best-performing stations, understand what made other stations fall short, and implement best practices across all stations to increase effectiveness across the board. Additionally, the model can be used in the future with updated data to obtain timely results and inform continuous improvements in LJI's transit-monitoring efforts to combat the societal ill of human trafficking.

Chapter 3

Modeling the Defensive Asylum Process in the United States Immigration Court System Using Queueing Theory

Section 3.1 Introduction

The United States has long been a country where people from all around the world immigrate, creating a place of unique diversity and culture. As immigration into the United States grew, immigration laws and policies also began to appear. As early as the 1880s, immigration laws restricted who could immigrate to the United States were passed, although the total number permitted was unrestricted. Eventually, the Immigration Act of 1924 established an annual limit to the number of immigrants who could enter the United States [247]. The current immigration policy limiting immigration is the Immigration and Nationality Act (INA) instated in 1952 with many amendments over the last 7 decades [248]. Despite existing limits on certain types of immigration into the United States, the overall demand remains greater than the allocated capacity [249, 250]. In the United States, nearly all immigration-related affairs are handled by one of six federal agencies with the exception of immigrants who enter illegally and manage to evade detection entirely [248]. Although the responsibility of immigration-related affairs is shared across six agencies, each handles specific functions. Therefore, the challenges and processes of each agency may differ, depending on the current demand across immigrants.

In recent years, the number of asylum-seeking immigrants has suddenly and significantly increased, mostly arriving at the United States Southwest border. Consequently, the demand across agencies handling asylum-related processes has dramatically increased, contributing to a growing backlog of cases and increasing wait times. A large increase in immigrants seeking asylum creates unique challenges because unlike other types of immigrants, such as Refugees, there is no limit to the number that can be annually admitted into the United States [251]. In fact, existing laws ensure that if an individual meets the qualifications for asylum, their case is heard in court [252]. While federal resources are designated to serve the immigration population, understanding the current resource utilization and the number of resources required to meet the demand proves challenging. Due to a policy passed on November 19, 2018, this problem was further aggravated, as many asylum seekers became entitled to an expedited process with the expectation that these cases would be adjudicated within 180 days [253].

Opportunities exits to improve the immigration court system where most of these asylum cases are processed. In this work, we focus on defensive asylum cases handled in the United States by the Executive Office for Immigration Review (EOIR), a branch of the Department of Justice. The EOIR is fashioned in what most would consider a typical civil legal setting, where individuals make claims or defend themselves against a claim and a judge decides how to adjudicate each case [248, 254]. The EOIR immigration system is very complex and therefore, a reasonable first step to improve the immigration court process is to understand how the current system functions. We address this problem through the mathematical modeling of the defensive asylum process using queueing theory. In Sections 3.2, 3.3 and 3.4, background information on the different types of immigrants, the defensive asylum process, and queueing theory are provided. Section 3.5 discusses related work at the intersection of immigration and analytics. Section 3.6 provides details on the data and data preprocessing steps. Section 3.7 presents the model representation and accompanying results. Finally, concluding thoughts and future work are discussed in Section 3.8.

Section 3.2
Background

Immigration is the international movement of a person from their country of origin to another (where they are neither native nor a legal citizen), with the intention to permanently reside there. In the United States, the INA defines an immigrant as a person seeking to become a Lawful Permanent Resident (LPR). LPRs are characterized as foreign nationals who have been permitted to work and live in the United States permanently [255]. There are many different types of immigration, including family-based immigration, employment-based immigration, and relief-based immigration. Both refugees and asylees fall under the relief-based immigration category, available specifically for vulnerable people(s) who are fleeing persecution or, for other reasons, unable to return to their country of origin safely [256]. Despite sharing characteristics of vulnerability and a need for protection, refugees and asylees have distinct differences which dictate how each is processed in the United States. In this work, we focus only on the processes involving asylum seekers.

3.2.1. Refugees

According to the American Immigration Council, a refugee is a person seeking legal admission into the United States based on an inability or unwillingness to return to their country of origin because of a " 'well-founded' fear of persecution due to their race, membership in a particular social group, political opinion, religion, or national origin" [256]. Refugees must apply for admission from *outside* of the United States, usually happening within a country outside of their country of origin, often referred to as a "transition country." A person's ability to be admitted into the United States as a refugee depends upon many factors, including their degree of risk, having a "well-founded fear" of persecution, belonging to a particular group of special concern according to the United States, or having family already established in the United States. The full legal definition of what constitutes a person as a refugee can be found in 101(a)(42) of the INA. While the total number of refugees that can be admitted each year has a general limit (referred to as the refugee admission ceiling), the exact yearly limits are decided by the President of the United States in consultation with Congress [257].

3.2.2. Asylum Seekers

Asylum seekers are individuals that meet the same general criteria required for a refugee but differ in the process by which they apply for relief. Unlike refugees who apply for relief from *outside* the United States, asylees must apply for relief from *within* the United States. More distinctly, refugees have their claims for relief vetted and approved for legal admission *prior* to arriving in the United States. In contrast, asylum-seeking immigrants must first arrive in the United States and then *subsequently* submit their claims for relief. There are two types of asylum an individual may pursue: i) affirmative asylum and ii) defensive asylum.

3.2.3. Types of Asylum

A defensive asylum-seeker is an individual who arrives (or enters) the United States without proper legal status and requests asylum as a mechanism to defend against being removed from the country. Individuals can claim defensive asylum at the port of entry or after being apprehended. There are two ways an individual may be apprehended: i) by a Customs and Border Patrol (CBP) officer within 100 miles of the United States border, or ii) by an Immigration Customs Enforcement (ICE) agent anywhere within the United States [258]. In both instances, the apprehended individual is then placed into what is called a removal proceeding. The defensive asylum process is handled by the Executive Office for Immigration Review (EOIR), a division of the Department of Justice (DOJ) [259]. An affirmative asylum seeker is an individual who has legally been admitted into the United States, is not currently involved in a removal proceeding and claims asylum. The affirmative asylum process is handled by the United States Citizenship and Immigration Services (USCIS), which is a division of the Department of Homeland Security (DHS) [259]. Although both types of asylum offer relief for vulnerable and persecuted individuals and require the applicants to be present in the United States, the processes for affirmative and defensive asylum differ. Affirmative asylum cases are handled by the USCIS in a non-courtroom setting, whereas defensive asylum cases are handled by the EOIR in a formal courtroom setting. Therefore, unless stated otherwise, henceforth, we refer to the EOIR as the immigration court system. While affirmative and defensive asylum cases are handled by different federal agencies, in some circumstances, an individual may interact with both agencies. For example, an individual may begin by submitting an affirmative asylum application to the USCIS, however, if denied and placed into a removal proceeding, they may subsequently submit a defensive asylum application to the immigration court system. In all instances, asylum applications must be submitted to the appropriate agency within one year of entry into the United States [259, 260]. In this work, we model the immigration court system and, therefore, only depict the defensive type of asylum.

Section 3.3 The Defensive Asylum Process

There are five main elements associated with a defensive asylum case: *Case Priority*, *Credible Fear Interview*, *Notice to Appear Issued (NTA)*, *Court Proceedings*, *Scheduled Calendar Hearings.* Not every case will experience all five elements, and some elements may appear more than once throughout the lifetime of a case. Although the journey a case takes through the immigration court system tends to be unique, there are a few constants; i) each case will have a unique case identifier and ii) eventually, each case will "leave" the court system, although when and how is case dependent. In what follows, we explain the five associated elements and how each impacts the flow of a defensive asylum case through the immigration court system. The elements are introduced in the order that they most often occur.

3.3.1. Case Priority

One element that impacts the journey of a case through the immigration court system is the case priority label, which has varying levels. For example, in January 2020 case priority was highest for unaccompanied alien children being held in government custody and detained (adult) individuals [261]. The policies determining which cases have priority over others change frequently, and therefore for simplicity, we separate cases into two classes: i) priority and ii) nonpriority. A case classified as priority generally receives quicker processing than its nonpriority counterpart. A memorandum from January 2020 specified that priority cases were to be completed within a 60-day period [261]. As priority cases are fast-tracked, the total time and path through the immigration court system may distinctly differ from that of nonpriority cases.

3.3.2. Credible Fear Interview

Generally, the defensive asylum process begins with an individual claiming asylum either at the port of entry or to the apprehending CBP or ICE agents. After an individual has made a claim for asylum, the CBP and ICE agents are required to administer a credible fear (or reasonable fear) interview. A credible fear interview is a screening process where an individual must establish a "significant possibility" that, at a formal court hearing, they could demonstrate they meet the requirements for asylum. If credible fear is established, the individual is issued a Notice to Appear and enters the immigration court system. If the initial interview finds the applicant insufficiently meets the requirements of credible fear, then the individual is either removed from the United States or, alternatively, the individual can appeal the findings, in which case an immigration judge will review the decision. If the immigration judge overturns the original decision and finds credible fear, the individual is issued a Notice to Appear and enters the immigration court system, otherwise, the individual is removed from the United States.

3.3.3. Notice To Appear

The Notice to Appear (NTA) document is issued to individuals by the USCIS with a list of specific immigration-related charges being brought against them. The NTA signifies the initialization of an immigration court case and indicates that the individual must appear before an immigration judge on the date listed on the NTA or on a future date to be determined. If an individual fails to appear at this initial hearing, they are subject to absentia removal, which means the individual has given up their rights to a formal hearing and can be deported [249, 262].

3.3.4. Court Proceedings

Court proceeding is the legal term used to refer to a formal process to describe if an individual is in violation of the law – in the immigration court context these laws pertain to immigration. All cases that enter the immigration court system will have at least one proceeding, typically more. Different types of proceedings include: expedited, administrative removal, deportation, and exclusion. Proceedings are the overarching process that facilitates a resolution of charges against an individual, concluding with a verdict, in which an individual is either absolved or found guilty. Often a proceeding requires multiple interactions with the immigration court at what is known as scheduled calendar hearings.

3.3.5. Scheduled Calendar Hearings

A scheduled calendar hearing is a formal stage in the immigration court process, where a case comes before a judge to handle details and take action regarding their proceeding. Two main types of calendar hearings occur: Master and Individual. Typically, the first case appearance before an immigration judge occurs at a Master hearing, and the final case appearance occurs at an Individual hearing. Generally, a case must be given a minimum of ten days between the NTA and the first scheduled Master hearing to allow the individual(s) to obtain representation and prepare. A Master hearing is unique in that there are a large group of individuals at any given scheduled slot waiting for the judge to call their case. Once a case has been called upon to appear before the judge, the interaction tends to be short, lasting less than 15 minutes. While the scope of a Master calendar hearing can include a myriad of things (Table 2.1), their purpose is generally to advise the asylee (and if the asylee is represented, their legal counsel) of details that may impact their preparation for the larger, civil court Individual hearing [263]. Individual hearings are longer in duration, typically around four hours, and are where evidence is provided, matters are contested, and a decision on the proceeding is made by the immigration judge. The scope of an Individual calendar hearing can include many additional elements (Table 3.1) with the main objective being to plead the case to the immigration judge, and reach a verdict on the proceeding; however, there are occasions where this type of hearing may end without a final decision being made [263].

Table 3.1: **Scope of Master and Individual Calendar Hearings.**[1]

Master Calendar Hearing
Advise the case member(s) of the right to an attorney or other representative at no expense to the government
Advise the case member(s) of the availability of free and low-cost legal service providers and provide the case member(s) with a list of such providers in the area where the hearing is being conducted
Advise the case member(s) of the right to present evidence
Advise the case member(s) of the right to examine and object to evidence and to cross-examine any witnesses presented by the Department of Homeland Security
Explain the charges and factual allegations contained in the Notice to Appear (Form I-862) to the case member(s) in non-technical language
Take pleadings
Identify and narrow the factual and legal issues
Set deadlines for filing applications for relief, briefs, motions, prehearing statements, exhibits, witness lists, and other documents
Provide certain warnings related to background and security investigations
Schedule hearings to adjudicate contested matters and applications for relief
Advise the case member(s) of the consequences of failing to appear at subsequent hearings
Advise the case member(s) of the right to appeal to the Board of Immigration Appeals
Individual Calendar Hearing
Making an opening statement
Raising any objections to the other party's evidence
Presenting witnesses and evidence on all issues
Cross-examining the opposing witnesses and objecting to testimony
Making a closing statement

Section 3.4

Queueing Theory

Queueing theory is the mathematical study of wait times, examining behavior, performance and how and why a queue forms. Queueing theory is used to model and understand existing queuing (wait) systems, or to design and optimize the performance of a new queue. The two main goals of queueing theory systems are i) to estimate the performance of the system, and ii) find the queue design that optimizes system performance [264, 265]. Queueing theory enables the measurement of queue characteristics such as the rate entities arrive to the system or queue, the time entities wait for service in a queue, and how quickly entities exit the system or queue. Queueing theory can be applied to model both basic queueing systems (e.g., a single queue), as well as complex systems that contain queues within queues, known as queueing networks.

[1] Source: Adapted from Immigration Court Practice Manual, Section 4.16, 2016.

3.4.1. Queueing Systems and Queueing Networks

At its simplest, a queueing system represents a stochastic process with a single queue and a single server. More complex queueing systems are comprised of multiple servers serving the same queue or a set of servers serving a multi-step process represented by a sequence of queues, as illustrated in Figure 3.1. Queueing systems can operate independently or in tandem to form a queueing network. Queueing networks are queueing systems where queues are connected by a routing network which represents the various paths an entity may take before leaving the system. The likelihood that an entity will traverse a particular route is represented by that route's transition probability. Queueing networks operate as stochastic processes that (often) use Markovian methods to capture underlying queue behavior. In the immigration court context, each hearing type (Master and Individual) has its own respective queue, and the path a case takes through the system is dependent on the case's complexity. For this reason, a queueing network is best suited to model the immigration court system.

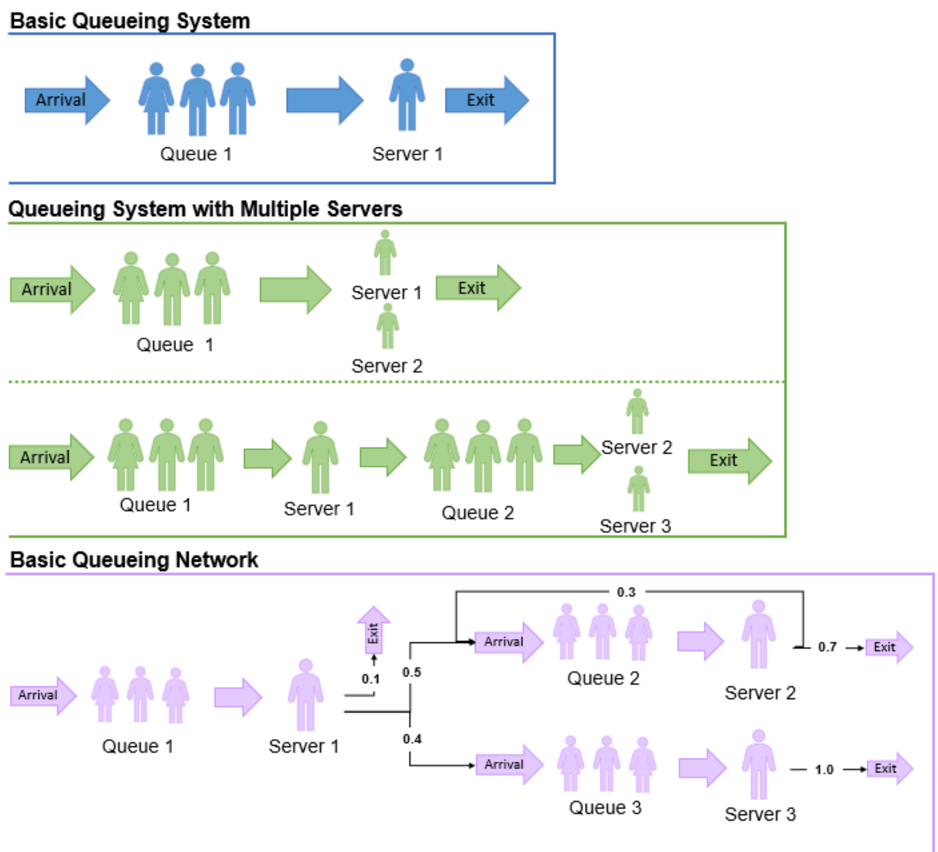

Figure 3.1: **Queueing System and Queueing Network.**

3.4.2. Characteristics of a Queue

Modeling a system using queueing theory requires the characterization of server and entity actions and behaviors within the system. Entities typically represent customers, or objects seeking service at a queue and a server represents those processing entities at a queue. An entity may take four actions regarding a queueing system: receive service, balk, renege, and jockey. Receiving service is the expected behavior – an entity arrives, waits in the queue until service, and then exits the queueing system. Alternatively, an entity may deviate from this standard behavior through balking, reneging, and jockeying. If an entity arrives at a queue but then decides to leave because they do not want to wait in a line, the action is called *balking*. For example, a customer arrives at a coffee shop and observing a queue decides not to wait, and leaves. When an entity is already queuing but decides to leave before being served, the action is called *reneging*. An example of reneging would be when a customer arrives at a coffee shop, gets in line, waits for ten minutes without service, then decides to leave rather than continue to wait. The term *jockeying* is used to describe when an entity enters one line, but then makes the decision to switch to a different line in hopes of reducing wait time. An example of jockeying would be if a customer arrives at a coffee shop and gets in line but then decides to switch to a different line in hopes of placing their order more quickly. A queueing system is defined by its queue discipline. *Queue discipline* refers to the order entities are served from a queue. While there are many possible queue disciplines to describe the server process, the most common designations can be found in Table 3.2. A common example of a queue discipline is First In First Out (FIFO), where the first entities to arrive are the first to be served.

3.4.3. Kendall Notation

Kendall Notation is a standardized, shorthand system consisting of six basic elements used to describe and classify a queue. Considering, simple queueing systems and complex queueing networks are comprised of the individual queues they contain, both can be represented using Kendall Notation. The long Kendall Notation takes the general form of $A/B/C/D/E/F$, where each letter represents a different element of the model. In the abbreviated Kendall Notation, the "$D/E/F$" elements are omitted when adhering to the default designated values. Table 3.2. shows what each letter represents as well as the different possible designations each element can take.

3.4.4. Stochastic Processes in Queueing

Queueing theory is based in probability theory, specifically stochastic processes. The arrival and service processes in queueing theory are commonly assumed to be stochastic in nature where the demands for service and interarrival times are treated as random variables. The memoryless property of Markovian processes is an attractive feature for modeling of arrival and service times in queueing theory. Arrival and service rates follow a Poisson process or, equivalently, interarrival and service times are exponentially distributed. In most cases Markovian assumptions are reasonable, allowing for the simplification and representation of otherwise mathematically complex queueing systems [264].

Table 3.2: **Kendall Notation.**[2]

Kendall Notation: A / B / C / D / E / F		
Symbol	**Name**	**Description**
A: Arrival Process		
M	Markovian	Poisson arrival rate with exponential interarrival times
M^x	Batch Markovian	Poisson arrival rate with x for number of arrivals per time arrival
D	Degenerate Distribution	A deterministic or fixed interarrival time
G	General Distribution	Independent interarrival times
B: Server Process		
M	Markovian	Exponential service time
M^y	Batch Markovian	Exponential service time with y for the number of served per time service
D	Degenerate Distribution	A deterministic or fixed service time
G	General Distribution	Independent service times
C: Number of Servers		
c	Fixed constant	Number of servers per queue, default value is 1
D: Queue Capacity		
d	Fixed constant	Capacity of queue, default value is ∞
E: Population Size		
e	Fixed constant	Population size being served by queue system, default value is ∞
F: Queueing Discipline		
FIFO	First in, first out	The first to arrive to the queue are the first to be served, the default value
LIFO	Last in, first out	The last to arrive are the first to be served
SIRO	Service in random order	The order of service is in random order, dependent on the order of arrival
PQ	Priority service	The order of service is dependent on an assigned priority

Section 3.5

Previous Work

The complex nature of immigration and its impact on civil society is a well-researched topic across varying domains such as law [266, 267], economics [268], sociology [269]. With the exception of one related study [270], our work is the first to look at the application of queueing theory to model the United States defensive asylum process.

3.5.1. Ongoing Analysis of the United States Immigration Court

An online tool exists that provides ongoing analysis and information on the United States immigration court system and case backlog. This tool is fittingly called the “Immigration Court Backlog Tool” [250], which has been developed and managed by the Transactional Records Access Clearinghouse (TRAC), an organization based out of Syracuse University. TRAC’s purpose is not only to provide information about the immigration court system, but more broadly, to provide the public with “comprehensive information about staffing, spending, and enforcement activities of the federal government" [250]. Starting in 1996, TRAC has developed several online sites that provide exploratory tools based on data they have collected, such as the aforementioned immigration court tool. The data used by TRAC originates either from existing publicly available resources or is obtained through the Freedom of Information Act (FOIA). The immigration court backlog tool uses data acquired from the DOJ through a FOIA request. As a result of numerous FOIA requests from TRAC over the years, the DOJ made the data publicly available in 2016 and continues to be updated semi-regularly. The TRAC tool is an excellent source to summarize the immigration court as a whole, however, a drawback of this tool is that the data is not readily available at a case level. The TRAC data is provided in the form of summary tables on an aggregated level, limiting the broader understanding of how the backlog has been created. As we are interested in modeling and understanding case-level queueing behavior we break down the data, analyze the process on a case-level view, and model it using queueing theory. Leveraging the data in this way can provide insights not only into what may have caused this backlog but also guide ways to improve the immigration court system moving forward.

3.5.2. Quantitative Studies of the United States Immigration Court System

Several studies measuring the United States immigration system exist. One in particular examines the impact participation in the EOIR's Legal Orientation Program (LOP) makes on the outcomes of immigration cases and the overall performance of the EOIR's court operations [271]. The LOP's intention is to provide information to detained immigrants about their rights in the immigration process. LOP's hope is that by providing this information to individuals, they will be able to make better-informed decisions throughout the duration of their case. The LOP uses representatives from non-profit organizations to provide this information as well as offer these individuals related services. The EOIR conducts analyses on asylum cases comparing individuals who participated in the LOP, with those that did not participate and every decade reports on the results [271, 272]. The LOP analysis report covers seven main areas, providing quantitative analysis and measuring the performance of each: length of stay at detention facilities, respondent representation rates, proceeding outcomes, proceeding length and case length, number of EOIR Hearings, the likelihood of proceeding and case completions, and adjournments and adjournment attribution. Although specifically focused on understanding the impacts the LOP has on an immigrant's time in the system, this report provides insight into the different elements of the immigration court and asylum processes. For this reason, this report is a useful source for understanding variation across cases, highlighting the complex nature and uniqueness of each case during its journey through the immigration court system. Such analysis is beneficial – the LOP is a part of the EOIR, and therefore, analysts are able to represent, understand and interpret the data with greater context, providing a more comprehensive analysis in comparison to those who are from outside the EOIR. Similarly, the drawback of the LOP report is that the analysis is commissioned by the EOIR themselves, and the report may be limited in explanations and exploration of certain concepts providing analysis on only things the EOIR wishes to share.

3.5.3. Queueing Theory and Immigration

To the best of our knowledge, at the time of writing, our work is the first to use queueing theory to model an immigration process in the United States immigration court system, with the notable exception of Rubio-Herrero [270]. In the aforementioned work, the author deploys queueing theory to model a precursory process of the immigration court by evaluating the performance of transition centers for a particular group of immigrants, Unaccompanied Alien Children (UAC) [270]. UAC are individuals under the age of 18 without the company of an adult and who have entered the United States illegally. UAC entering the immigration court system undergo a slightly different process from that of the general population and must be placed into a state-licensed shelter, referred to as a transitional center, within the first 72 hours of apprehension. In 2015, an influx of UAC entering the immigration court system created an overflow at the transitional centers and consequently, created difficulty in adhering to the 72-hour rule. Seeking improvements to capacity utilization, Rubio-Herrero utilized queueing theory to analyze the performance of these transition centers by considering environmental factors such as capacity, nature of arrivals into the system, and service time. Rubio-Herrero created a straightforward model of this process through the use of a well-known queueing model, $M^x / M / c$. Using this model the author explored different scenarios, varying the environmental factors. Results from this revealed how the capacity utilization of a transition center changes depending on the arrival and service times. This work demonstrates the potential of applying queueing theory to immigration-related problems. Our work extends the use of queueing theory in the immigration setting by looking at the broader immigration court setting.

Section 3.6

Data

The EOIR case data (CASE) represent cases from the United States immigration court system. The CASE data have been made publicly available since 2016 due to multiple FOIA requests by TRAC over the years [250]. This data is provided in raw files originating from the EOIR's electronic database, downloadable from the DOJ's website in the form of a ZIP file [273]. The CASE dataset contains information on numerous aspects of each case in the EOIR immigration court system from 1950 to the present day and is updated semi-regularly. As of February 2, 2020, the dataset consisted of 19 tables containing data on cases, 78 lookup (reference) tables, 1 schema table, and the "data code key" (Figure 3.2). Our analysis focuses on the sub-population of defensive asylum cases, and therefore we were able to reduce the number of tables, features, and data necessary for our model.

#	Field Name	Business Name	Table	Definition	Data Type	Data Sample
1	IDNCASE	Case Table Index	A_tblCase	The primary key for A_tblCase	int	6125515
2	ALIEN_CITY	Alien Address City	A_tblCase	Alien address city	varchar(15)	SACRAMENTO
3	ALIEN_STATE	Alien Address State	A_tblCase	Alien address state	varchar(2)	CA
4	ALIEN_ZIPCODE	Alien Address Zip Code	A_tblCase	Alien address zipcode	varchar(10)	95842
5	NAT	Alien Nationality	A_tblCase	Nationality of the alien	varchar(2)	ZI
6	LANG	Alien Language	A_tblCase	Primary language spoken by	varchar(3)	ZUL
7	CUSTODY	Alien Custody Status	A_tblCase	Current custody status of the	varchar(1)	R
8	SITE_TYPE	Hearing Notification Type (DHS)	A_tblCase	Hearing notice issued IN-PERSON or MAILED to the alien	varchar(12)	M
9	E_28_DATE	Block 3 update data	A_tblCase	Legacy field- no longer used	varchar(2)	9912301332
12	ATTY_NBR	Number of Attys	A_tblCase	Legacy field. Refer to tbl_RepsAssigned data	varchar(3)	1
13	CASE_TYPE	Case Type	A_tblCase	Type of case	varchar(3)	RMV
14	UPDATE_SITE	Current Base City Code	A_tblCase	The current base city code of the case.Same as base city code on current proceeding.	varchar(3)	YOT

Figure 3.2: **Example of EOIR Case Data Code Key.**

Six fields are required to determine cases that are both i) defensive asylum and ii) within our modeling scope: *Application Type*, *Appeals Filed*, *Rider or Lead Case*, *Case Status*, *Schedule Hearing Types*, and *Case Priority*. The case type must be a defensive asylum case, where the individual has entered into the EOIR immigration court system as a result of being apprehended and submitted an asylum application (at some point during their case). Case appeals are handled outside of the EOIR and are excluded from our data. Sometimes cases are connected by a Lead or Rider flag; a Lead case is considered to be the primary case and the Riders are associated cases. The Rider cases experience the same process through the immigration system and therefore to avoid double counting these instances, we keep only the Lead cases. We wish to capture the entire duration of a case from start to finish, and for that reason, we only include cases that have reached completion. The vast majority of cases only consist of two types of hearings: Master and Individual. In an effort to simplify an already complex system, we consider only these hearings – cases with other types of hearings are removed. Lastly, per policy [261], detained cases are prioritized in scheduling, experiencing a different time through the system, and therefore an important feature to include when modeling.

3.6.1. Characteristics of the 10-Year Dataset

We consider 10 years of data, capturing all cases that have exited between 2010 - 2019. Figure 3.3 depicts the arrival year for cases on the x-axis and the total number of cases arriving in the system each day on the y-axis. A closer examination of this plot shows the vast majority of cases exiting between 2010 - 2019 entered the system between 2000 - 2019. Figure 3.4, shows the average daily number of arrivals to the system by year for cases that exited between 2010 - 2019.

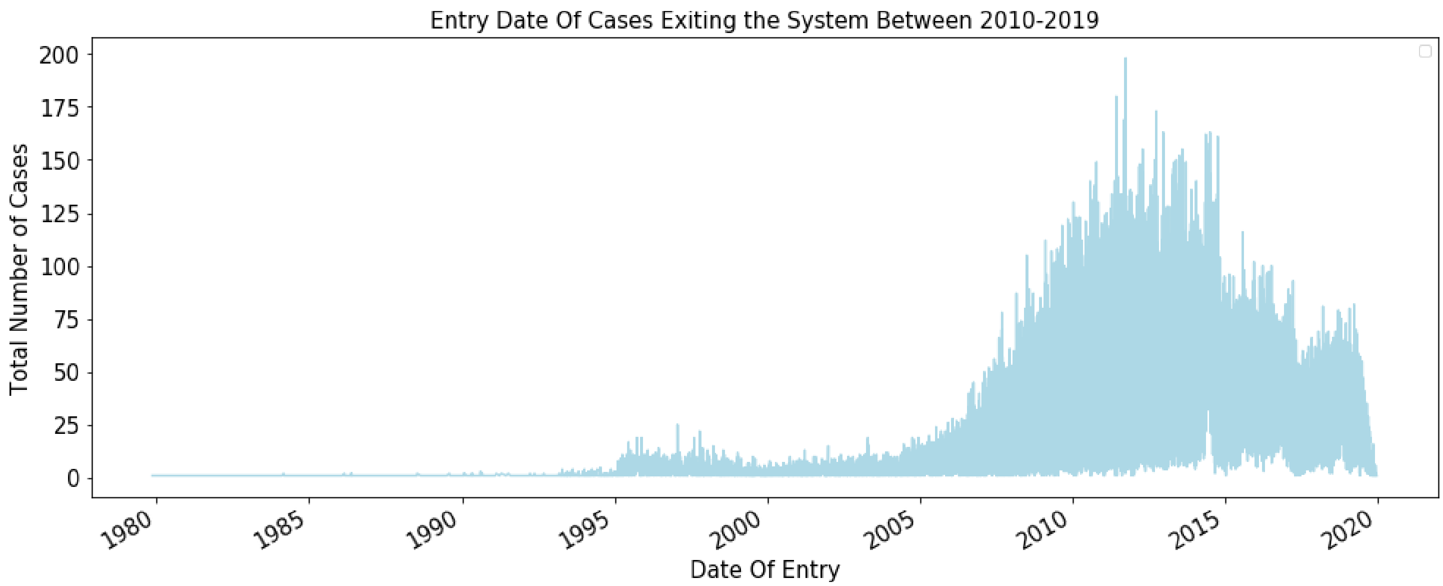

Figure 3.3: **Date of Entry for Cases in Datasets.**

Figure 3.5 shows the average total time a case spends in the system (sojourn time) plotted by the year they entered. For example, the average sojourn time for cases that *entered* the system in 2013 was approximately 538 days (almost a year and a half). Considering we only include cases that have *exited* (reached completion) between 2010 - 2019, the presented plots show a decreasing pattern, as cases still in progress are not displayed (their total sojourn time is not yet known).

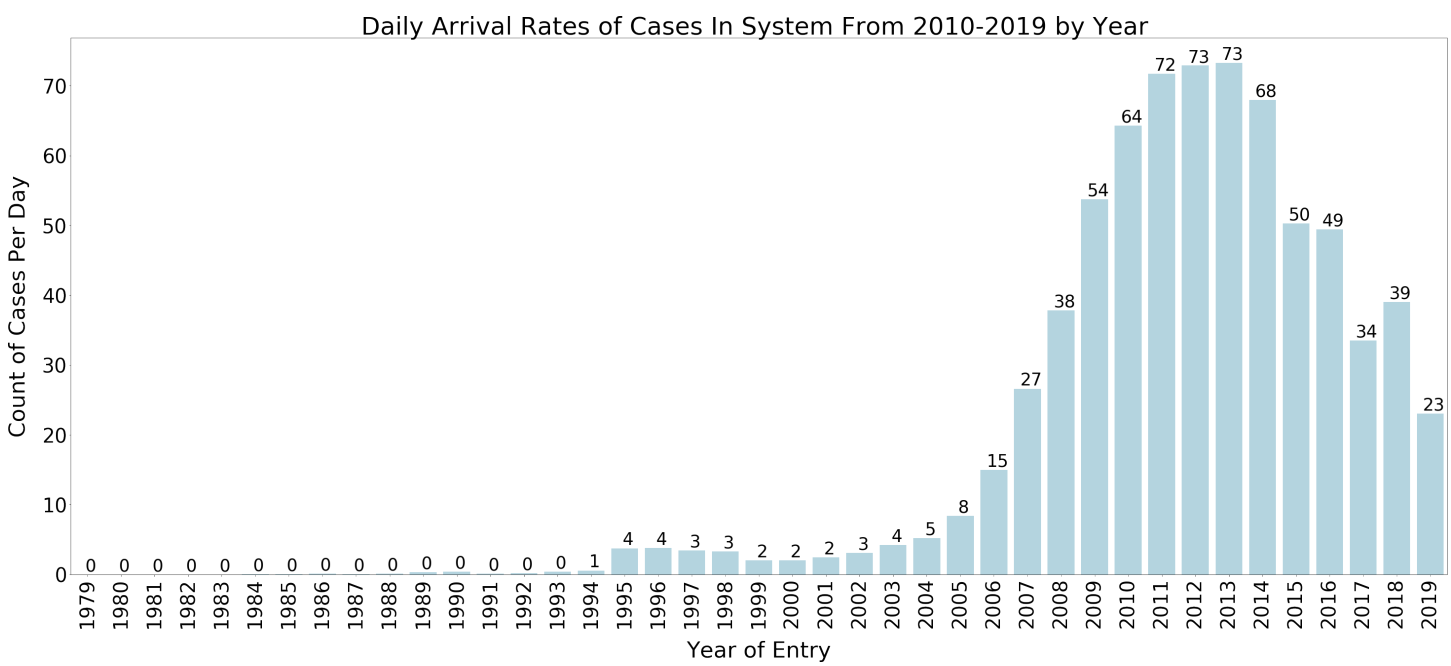

Figure 3.4: **Daily Arrivals to System by Year.**

3.6.2. Elements of Cases in the Dataset

Proceedings and scheduled hearings are the two main elements we consider in modeling the immigration court system. In Figure 3.6 we see that by far, most cases will have only one proceeding during their time in the immigration court system. In general, cases with more proceedings tend to have longer sojourn times (total time in the system).

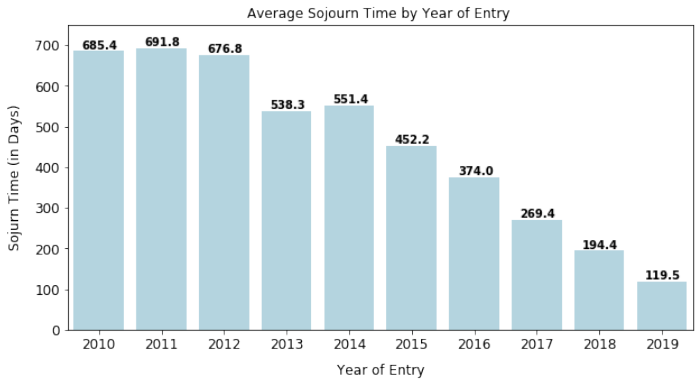

Figure 3.5: **Average Sojourn Time by Year of Entry.**

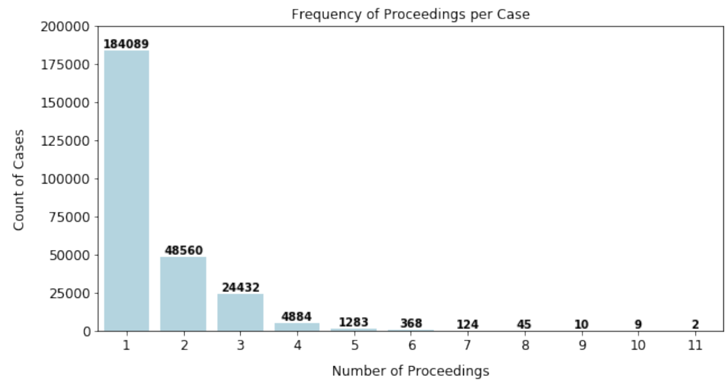

Figure 3.6: **Frequency of Proceedings per Case.**

Another element that impacts the sojourn time of a case is the number of scheduled hearings. Figure 3.7 depicts the total number of scheduled hearings per case with the vast majority of cases having more than one. Considering the sojourn time of a case is dependent on the number of hearings, appropriately representing the hearings and when each occurs is important in the modeling process. Figure 3.8 depicts the hearing order (x-axis), and the number of Master and Individual hearings observed for each hearing order (y-axis). The first scheduled hearing for nearly all cases was a Master hearing, however, as the sequence of a case progresses Individual hearings become more common.

Figure 3.9 depicts a sample of 30 cases, where Master hearings are depicted by "M" (light blue bars), and Individual hearings are depicted by "I" (dark blue bars). Master hearings most commonly occur in the first few occurrences for a case and Individual hearings are most common towards the end of a case. Individual scheduled hearings are generally where a judge provides a verdict. This figure showcases the variability and idiosyncrasies of case paths through the system. The demonstrated variability in paths through the system provides further motivation to model the immigration court using a queueing network, where transition probabilities can represent the likelihood of having a Master or Individual hearing.

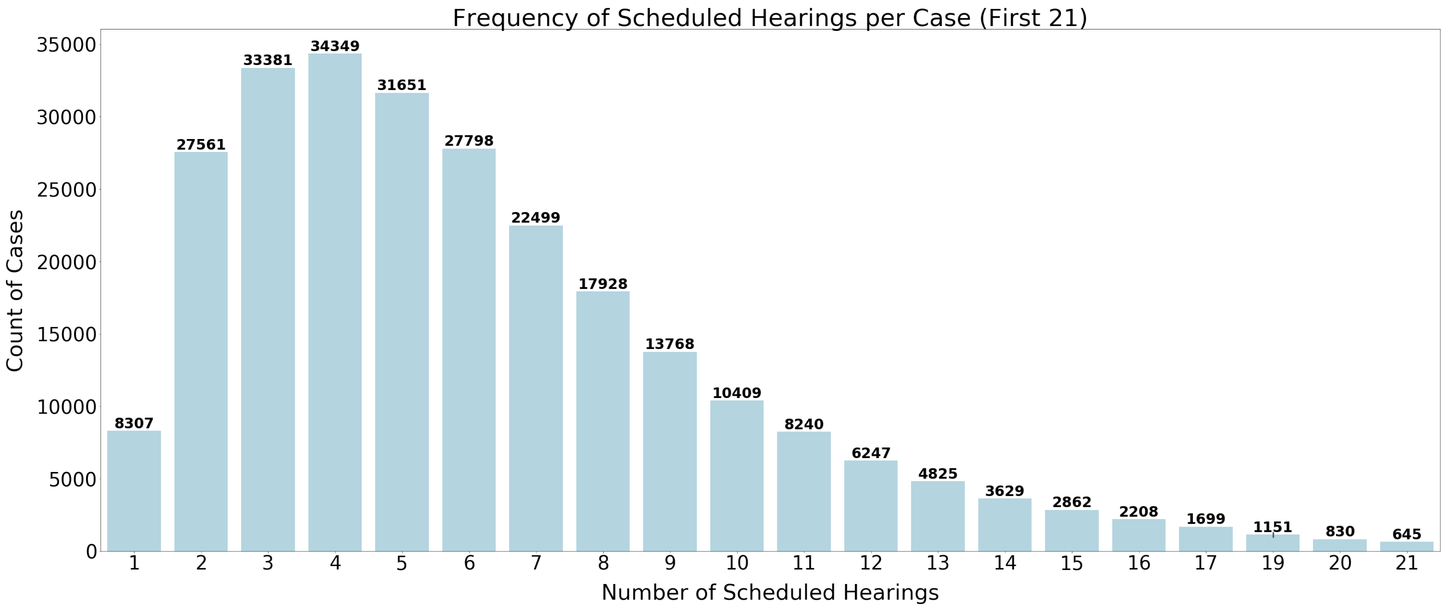

Figure 3.7: **Frequency of Scheduled Hearings per Case.**

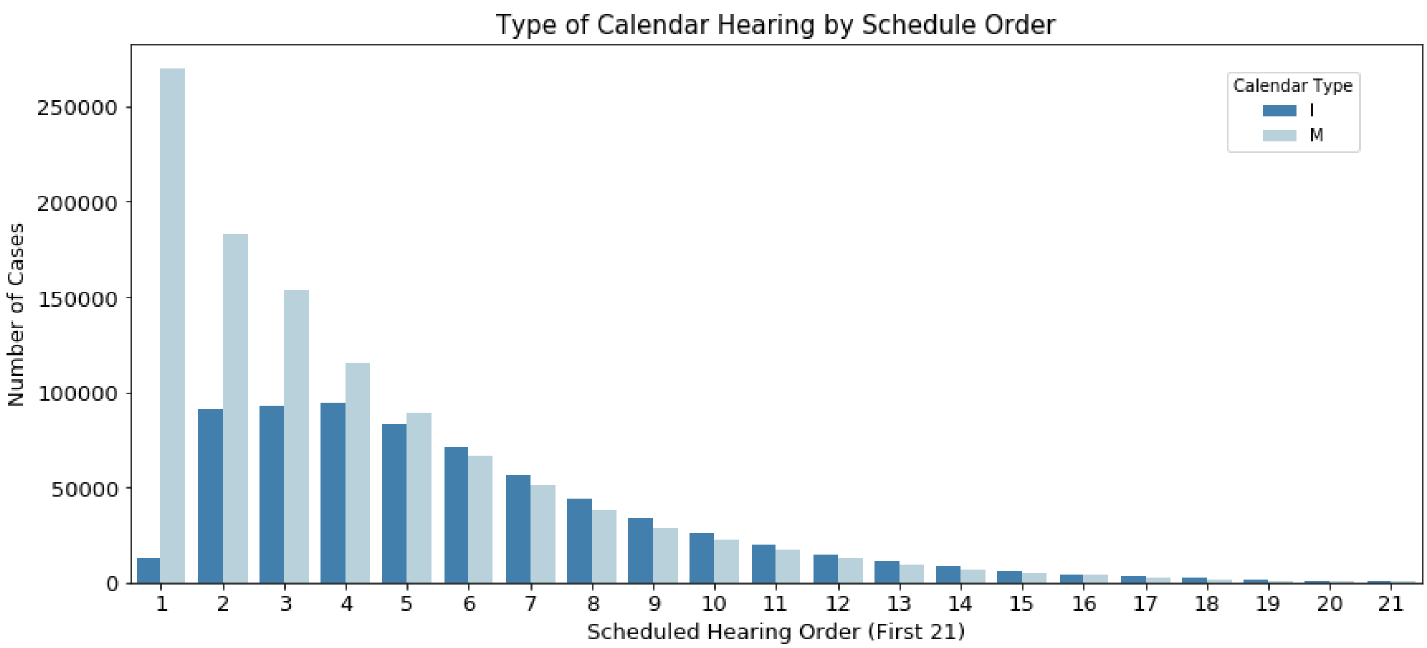

Figure 3.8: **Type of Calendar Hearing by Scheduled Hearing Order.**

Section 3.7

Our Model

The immigration court system is a process where cases are handled in some fashion– queueing theory allows this process to be analyzed by breaking the process down into elements. The immigration court contains multiple hearings (Master and Individual) that can be represented by two separate queues. Stochastic processes can capture how cases route between these queues and enter and exit the system. Arrival and service rates encapsulate system

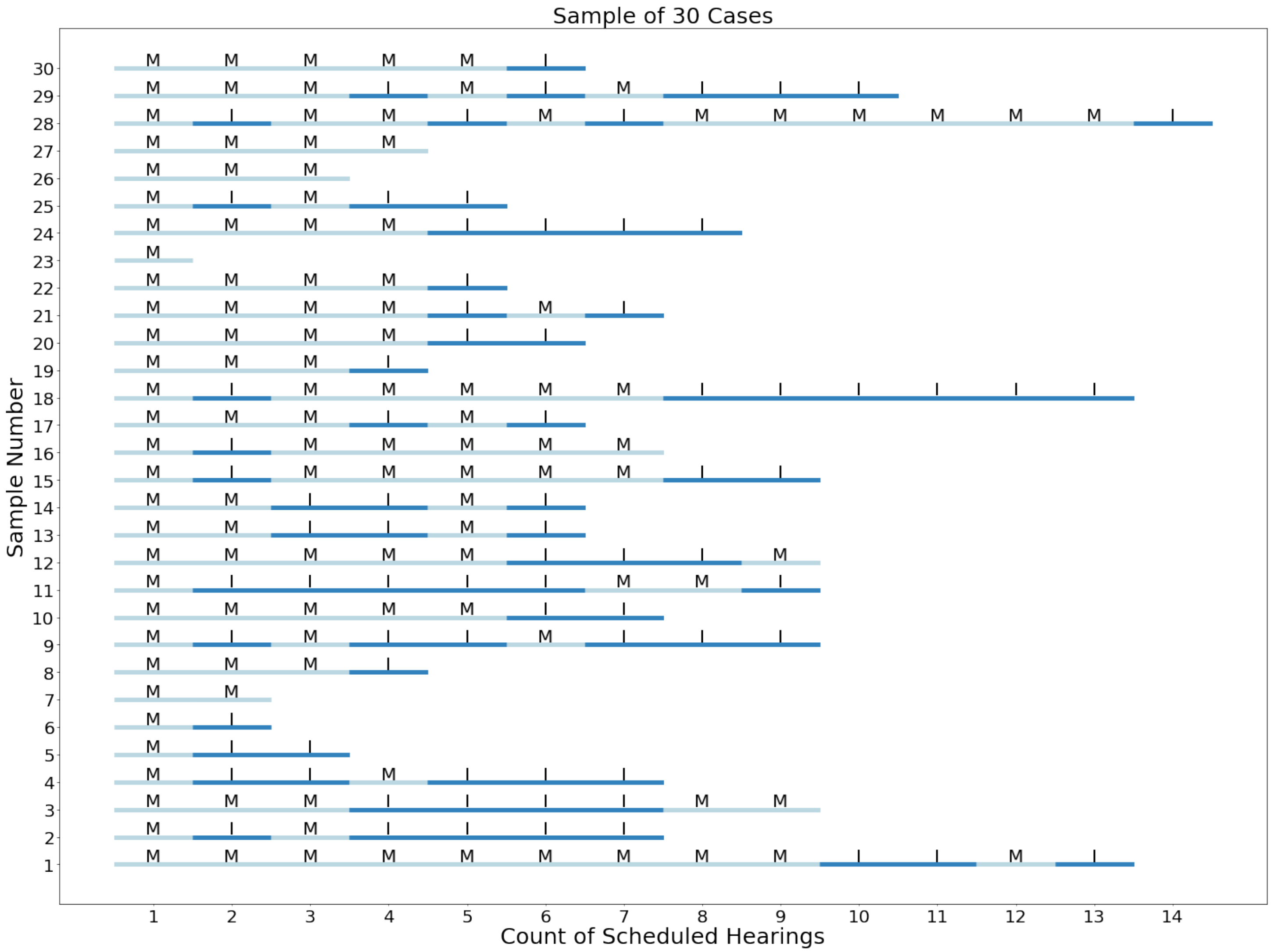

Figure 3.9: **A Sample of Scheduled Hearing Patterns for 30 Cases in the System.**

capacity and traffic; how quickly cases arrive and how quickly cases are served. In queueing theory, the wait, service, and sojourn times summarize the behavior of cases through a system helping to identify bottlenecks. We now present our model formulation and values used to produce a queueing framework and visual representation of the United States immigration court defensive asylum process.

3.7.1. Our Model Formulation

The queueing model best suited for modeling the defensive asylum immigration court process is a multi-class open queueing network with two queues, where both queues have batch arrival and service processes, one server, infinite capacity, and a priority service queue discipline. Cases in the immigration court system with a priority label are handled differently than nonpriority cases. To capture this effect, we classify our model as a multi-class system where one class represents the priority cases and the other nonpriority. Master and Individual hearings serve different purposes and therefore are represented as separate queues. On a daily basis, multiple cases arrive and exit the system and queues, thus, both arrivals and services are represented as batch processes. All cases exiting a queue are served in batches once a day and for this reason, each queue has a single server. As there is currently no limit to the number of asylum cases that can enter the immigration court system our model has infinite capacity. The queue discipline for each queue is priority service. Priority service in this context means cases labeled as priority (detained), are served before nonpriority (undetained); however, cases with equal priority are served following FIFO. In Kendall Notation each queue can be represented as $M^x/\ M^y\ /1/\infty/PS$. The presented model holds three assumptions: i) interarrival times are exponentially distributed ii) service times are exponentially distributed and iii) the system is memoryless, as described below. Although a case may have two hearings of the same type back-to-back, for simplicity, consecutive occurrences of a hearing type are represented as a single unit of service. Master hearings tend to occur before Individual hearings and therefore are referred to as queues one and two, respectively. Together, these two queues form an open queueing network where cases freely enter and exit the system. The routing network of the system is comprised of seven unique transition routes (described in Table 3.5), each with its own probability.

Table 3.3: **Seven Transition Routes.**

Route	Transcription Description
1	After Entering the System: Enter Queue 1 First
2	After Entering the System: Exit Before Ever Visiting a Queue
3	After Entering the System: Enter Queue 2 First
4	After Service from Queue 1: Enter Queue 2
5	After Service from Queue 1: Exit the System
6	After Service from Queue 2: Exit the System
7	After Service from Queue 2: Enter Queue 1

Using the data described in Section 3.6, transition probabilities are estimated and the average observed values for i) arrival rates, ii) wait times, and iii) service times are obtained for each queue and described further in Section 3.7.2. The visual framework of our model can be seen in Figure 3.10.

3.7.2. Our Model Values

Our queueing model requires six sets of values estimated from the data: *Arrival Rates*, *Service Rates*, *Wait Times*, *Service Times* (sojourn time in queue), *Total Sojourn Time* (in system) and *Transition Probabilities*. In this section, we discuss how these values were calculated. An important characteristic of our model is that the $M^x/\ M\ /1/\infty$ queueing model is a specific case of time-homogeneous, continuous-time Markov chain. This means both interarrival and service times are exponentially distributed where the transitions between states do not depend on time and are thus memoryless [264, 274]. The memoryless property implies that each transition probability

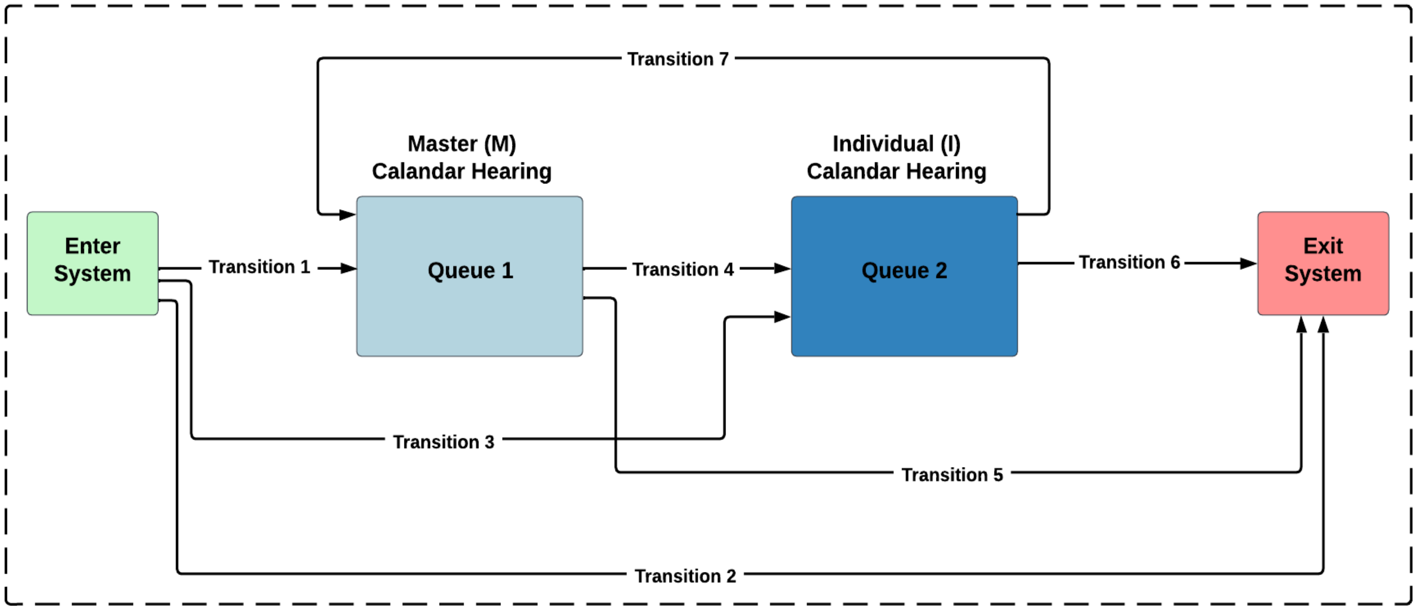

Figure 3.10: **Queueing Model Framework.**

Table 3.4: **Transition Paths.**

Transition	Path A	Path B
4	1,4	3,7,4
5	1,5	3,7,5
6	3,6	1,4,6
7	3,7	1,4,7

is independent of past events so calculating the transition probabilities does not require exhaustively finding every path a case takes through each queue, but rather, limits the search to only finding the first time a case passes through each queue. Limiting the data to include only paths that pass through each queue once avoids overestimating the probability of each edge. The probabilities for Transitions 1 and 3 were calculated using the percentage of cases whose first scheduled hearing after arriving into the system were *"M"* and *"I"*, respectively. The probability of Transition 2 was calculated by taking the percentage of cases who exited the system without a single hearing. Transitions 4 through 7 were not as straightforward to calculate considering for each of these transitions, there were two unique paths: A and B. Both of these paths first passed through a queue(s) before exiting via these transitions (Table 3.4). We use Figure 3.11 to illustrate these different paths.

Transitions 4 and 5 first pass through Queue 1 prior to i) entering into Queue 2 or ii) exiting the system, respectively. Therefore, while Transition 1 (in yellow) is a path independent of any other path, a case enters Queue 1 directly without passing through another queue; the transition probability for Transitions 4 and 5 is dependent on Transition 1 and is represented by Path A. Alternatively, Transitions 4 and 5 can be dependent

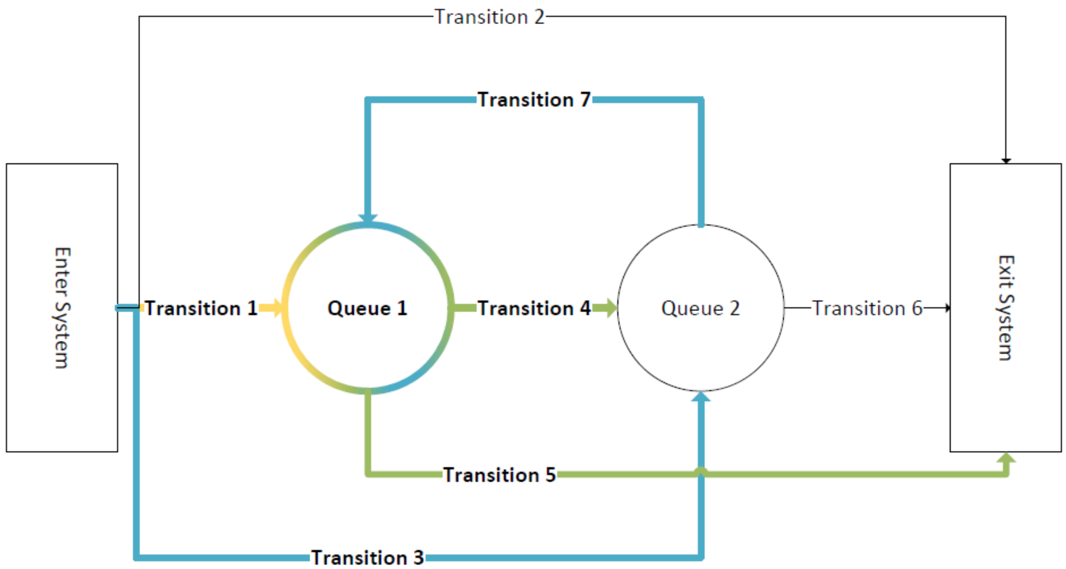

Figure 3.11: **Example of Paths A, B for Queue 1.**

Table 3.5: **Transition Probabilities.**

Transition	Priority	Nonpriority
1	88.83%	99.56%
2	0.0010%	0.23%
3	11.17%	0.21%
4	94.06%	81.72%
5	5.94%	18.28%
6	34.82%	63.67%
7	65.18%	36.33%

on first passing through Queue 2 via Transitions 3 and 7 and are represented by Path B. The calculations for Transitions 6 and 7 follow a similar logic, where Path A represents these transitions first passing through Queue 2 and Path B represents these transitions first passing through Queue 1 via Transitions 1 and 4 prior to either i) exiting the system or ii) entering into Queue 1. For Transitions 4 through 7, the probabilities of both Path A and B must be considered to obtain accurate transition probabilities. For example, the percentage of cases passing through Queue 1 via Path A and B and entering into Queue 2 is the probability of Transition 4. Transitions 5, 6, and 7 follow the same logic. All of the transitions probabilities used in our model can be seen in Table 3.5.

The arrival rates, service rates, service times and wait times for Queue 1 and Queue 2 are also dependent on these different paths (Table 3.6), following similar logic with one big distinction; these values are not probabilities. Arrival and service rates are expressed in the number of cases per day and wait and service times are expressed in days.

These values are calculated using the weighted average of cases arriving/being served from Path A plus the

Table 3.6: **Paths Entering Queue 1 and Queue 2.**

Queue	Path A	Path B
1	1	3,7
2	3	1,4

Table 3.7: **Arrival Rates Into Queues (Cases Per Day).**

Queue	Priority	Nonpriority
1	12.12	19.77
2	12.19	16.91

weighted average of cases arriving/being served from Path B, as shown in equations 3.1 (or alternatively 3.2).

$$\text{Arrival to Queue N: } \frac{A}{N} * \text{Rate A} + \frac{B}{N} * \text{Rate B} \tag{3.1}$$

$$\text{Arrival to Queue N: } \frac{A}{N} * \text{Rate A} + 1 - \frac{A}{N} * \text{Rate B} \tag{3.2}$$

Where:

N = The Number of Cases in Queue N

A = Number of Cases arriving/being served to Queue N from Path A

B = Number of cases arriving/being served to Queue N from Path B

Arrival rates represent the average daily number of cases entering each queue and are displayed in Table 3.7. Service rates for each queue represent the average daily number of cases on the last (consecutive) hearing of that type (Master or Individual), exiting the queue and are displayed in Table 3.8.

Considering, consecutive hearings of the same type are represented as a single unit of service, the average service times for each queue are expressed by the summation of service times for these hearings and are displayed in Table 3.9. The wait time for each queue represents the average number of days between a case entering a queue and exiting that same queue, as seen in Table 3.10.

Table 3.8: **Service Rate Leaving Queues (Cases Per Day).**

Queue	Priority	Nonpriority
1	34.72	44.90
2	17.96	21.71

Table 3.9: **Service Times (in Days).**

Queue	Priority	Nonpriority
1	219.75	242.83
2	46.64	207.51

Table 3.10: **Wait Times (in Days).**

Queue	Priority	Nonpriority
1	83.17	111.88
2	395.55	765.81

The final system values estimated are the total sojourn time for each case in the immigration court system, and case arrival rates to the system. The total sojourn time represents the average number of days between when a case first enters the system and completes its last hearing in the system, Table 3.11 shows these values. The system arrival rates represent the average daily number of new cases entering the system and are shown in Table 3.12. All values are calculated for both Priority and Nonprority cases and depicted within the queueing framework in Figure 3.12.

Table 3.11: **Total Sojourn Time (in Days).**

Priority	Nonpriority
371.98	1,242.78

Table 3.12: **Arrival Rate to System (in Days).**

Priority	Nonpriority
14.80	19.84

3.7.3. Interesting Characteristics of Queueing Model

Examining Figure 3.12 more closely affirms known characteristics of the defensive asylum process as well as reveals some interesting insights. Taking a closer look at the transition probabilities it can be seen that cases can enter the system for the first time via Transition 1, 2 or 3. Transition 1 represents cases whose first scheduled hearing is a Master calendar (Queue 1) hearing and appears to be the most common first transition into the system. This observation aligns with the intent of Master calendar hearings, which is to provide time for asylees to address an array of preliminary details for their case, often in preparation for an Individual calendar hearing (Queue 2). The chances of both priority or nonpriority cases entering the system, but then leaving (Transition 2) prior to receiving service at either queue is very low. This observation is not shocking, given the vulnerable nature of asylum seekers and the level of risk it often takes these individuals to be able to apply for asylum in the United States in the first place [275]. However, we hypothesize this value might be larger for general removal cases. Transition 3

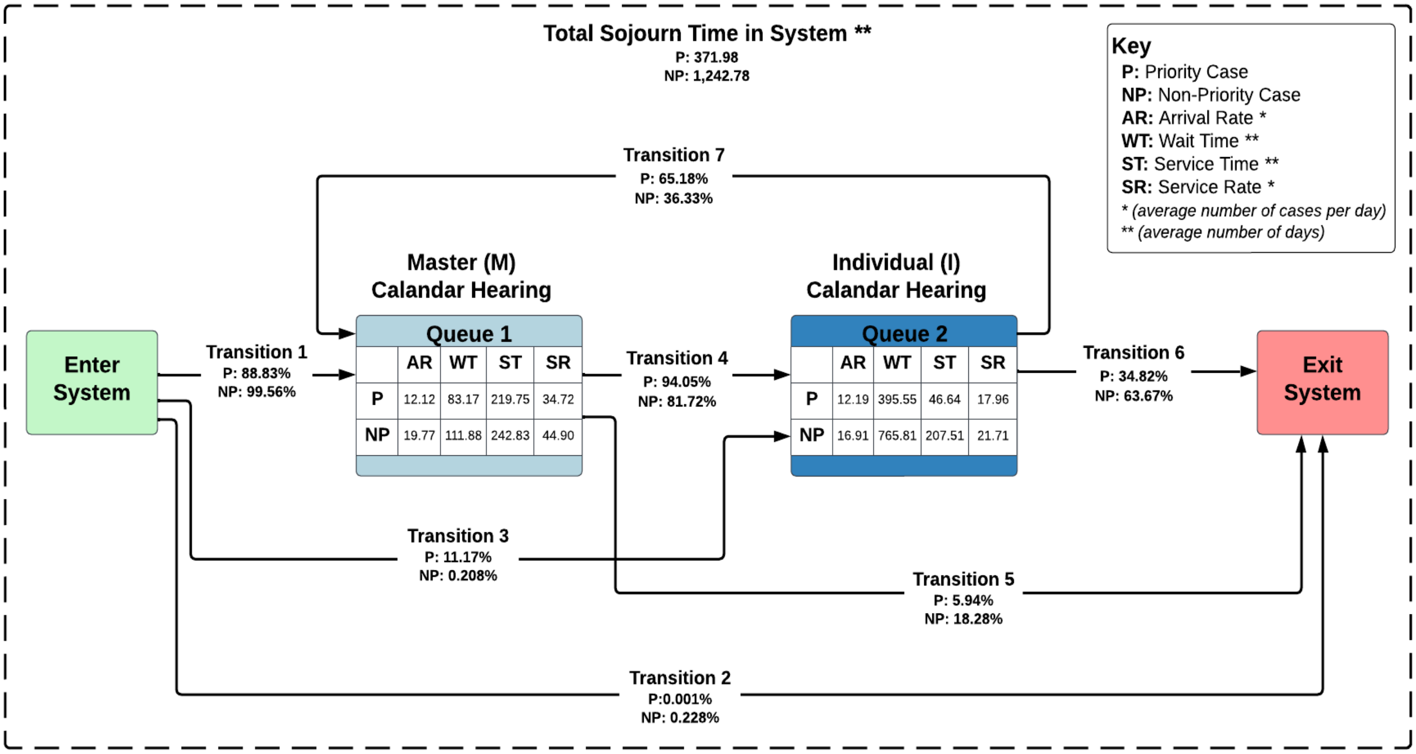

Figure 3.12: **Final Visualization of Queueing Model.**

represents cases whose first scheduled hearing is an Individual calendar hearing. Comparing the two classes reveals that priority cases are more likely to skip the Master hearing and instead go directly to an Individual hearing, indicating priority cases may have unique circumstances, such as policy or processes influencing this difference in behavior.

The majority of both priority and nonpriority cases proceed to an Individual hearing after receiving service at Queue 1 via Transition 4, however, it should be noted, nonpriority cases are more likely to exit the system after receiving service at Queue 1 (Transition 5). Comparing the behavior of priority and nonpriority cases after receiving service in Queue 2, reveals an interesting difference, each class transitions in nearly the exact opposite way. Priority cases are more likely to transition back into Queue 1 (Transition 7), whereas, with almost the same probability, nonpriority cases are more likely to exit the system (Transition 6).

The average daily arrival rate of cases into each queue is nearly the same for priority cases, and only slightly different for nonpriority cases. Comparing the arrival rates between the two classes of cases indicates a slightly higher rate for nonpriority, which is to be expected, considering nonpriority cases account for a larger percentage of all cases entering the system (as seen in the average number of arrivals to the system). Aside from Transitions 6 and 7, as noted above, the transitions of priority and nonpriority cases behave in a fairly similar manner.

Although priority and nonpriority cases transition similarly throughout the system, these cases tend to differ in the total time spent in the system. The total sojourn time for priority cases lasts on average around one year, whereas nonpriority cases stay in the system for over three times that amount. One potential contributing factor to this may be that priority cases are served faster than nonpriority cases in Queue 2, where priority cases are completed in a little over a month and nonpriority on average take nearly six months. Wait times for Queue 2, however, are much longer than that of Queue 1. Considering Individual hearings require a larger amount of resources, each case taking approximately four hours, a larger wait time for service is to be expected. Master hearings (Queue 1) are shorter in length, and therefore judges tend to process more cases per day, consequently, a lower wait time and higher service rate for Queue 1 are observed.

Section 3.8

Conclusion and Future Work

Through the application of queueing theory, we developed an framework and representation of the defensive asylum immigration court process, estimating arrival rates, service rates, wait times, and sojourn times of both priority and nonpriority cases in the system. The presented model provides an opportunity to better understand the construction and flow of defensive asylum cases through the immigration court system. As demonstrated throughout this work, the immigration court system is complex; and thus modeling the defensive asylum process is no less so. We developed a baseline model of this system represented as a Multi-Class Open Queueing Network comprised of two queues following $M^x/\ M^y\ /1/\infty/PS$.

Future work can leverage this model to explore avenues for improved queue design and the optimization of the immigration court system processes. The result of this work provides a first step towards helping achieve this goal. The implications of our work go beyond simply improving a system: its success can positively affect asylum seekers as well. The mathematical modeling and implementation of new systems will contribute to societal benefits by allowing for the ability to analyze the system for more fair policies and ways to lower the overall time that asylum cases spend in the system. The intention of this and future work, are not to be the final solution to the problem, but rather to be a tool used to help inform decision makers, such as those deciding policy related to the processing of defensive asylum cases.

There are many possibilities for future research and further development of this work. As previously expressed, this work is a simpler model of the complex system. For example, an assumption made was that all consecutive scheduled hearings in a queue represent a single unit of service and that these service times are exponentially distributed. In reality, the service times are in phases and therefore we would need to explore the modeling of these aspects. For instance, through the use of a phase-type distribution, such as Erlang or Coxian, these service times can be made to look close to Markovian and may in turn provide fairly good estimates.

Another extension of this work could be to expand the elements modeled in our system, such as including more than two queues and simulating the system. Through simulation, the servers and queues of complex systems can be determined and used to find an analytical solution. A simulated model or analytical solution would provide additional credibility needed for the deployment of a tool that could be used by decision makers to inform policy. This type of tool would allow the user to vary different queueing values, such as service time and the number of judges (servers) to evaluate system impacts. Based on this research we now consider some future extensions. From our model we can explore four directions: i) how cases should be scheduled, ii) how much time should be allocated for each scheduled hearing, iii) how does changing the way priority classes are assigned impact the system, and iv) how many judges are needed to meet the demand for different levels of traffic intensity.

Chapter 4

Modeling the United States Immigration Court Using Discrete Event Simulation

Section 4.1 Introduction

As of January 2023, over 2 million cases seek relief through the United States immigration court system, a volume growing rapidly with unprecedented influxes of migrants at the southern border [276]. Many immigration court cases involve asylum seekers, and the decision of a case can significantly impact an individual's life trajectory. All immigrants, including those undocumented, possess basic rights under the United States Constitution. The overwhelming demand from migrants coupled with limited governmental resources has greatly impacted the ability of the immigration court to adjudicate cases in a timely manner. Thus, improving case processing is of great interest to both immigrants and the United States government. As of January 2023, the average time for a case to complete was 1,016 days [276]. Government agencies and the media have stressed the need to address the growing backlog of immigration court cases [277,278], and it is a well-documented area of interest for the current administration [279].

Despite this attention, research to develop and evaluate solutions to the backlog appears limited [280]. This gap presents an opportunity to employ analytics to improve the immigration court system. The immigration court is a complex system with arrival rates, service rates, and scheduling rules that can vary by day and case. An individual's progress through the system depends on the unique characteristics of the case and system idiosyncrasies. Discrete event simulation (DES) is an established method across various domains such as healthcare [281, 282], manufacturing [283], and transportation [284], that provides flexibility in modeling complex decision logic, system structures, and rules. We explore the use of DES in immigration court systems, focusing on the United States.

We establish a baseline DES model by simulating nearly 1 million hearings across multiple judges within the New York City immigration court. We validate our baseline model over 15 years of historical data. Data Science (DS) techniques are used to determine daily arrival rates, service capacities, and entity attributes. We use historical wait times to assign a relative priority to cases and infer a processing order for each hearing using a "ticketing" approach. This work uses DES and DS to model the system and explore ways to improve its throughput. Our baseline model creates an analytical representation that can be expanded to evaluate, inform, and improve immigration court policies seeking ways to alleviate the United States immigration court backlog.

The chapter proceeds as follows. Section 4.2 provides context for how our work contributes to using Applied Analytics to address problems within immigration operations. Section 4.3 provides additional details on our modeling context and data. In Section 4.4 we provide the baseline DES model formation and discuss model validation. Section 4.5 provides a sensitivity analysis of our model by varying arrival and service rates. Section 4.6 concludes the presented work and points to future improvements and applications.

Section 4.2

Background

Analytics has been applied to address several immigration-related problems in the United States. Ahani et al. [3] use machine learning and integer optimization to develop a novel decision-making tool to assist the placement of refugees within the United States. The aforementioned model is extended to account for risk in the placement of refugee families [4]. The authors show that risk-averse optimization models can reduce risk, while maximizing expected outcomes. Considering varying environmental factors, Rubio-Herrero [270] apply queueing theory to assess the performance of transition centers serving the vulnerable immigration population of Unaccompanied Alien Children. De Azevedo Drummond [285] develop a bayesian decision model to examine operations at the United States port-of-entries, prescribing an optimal policy for the screening and detection of human trafficking victims. Irvine et al. [286] examine the impacts of COVID-19 on the United States Immigration and Customs Enforcement detention facilities, seeking to understand the rate of spread of COVID-19 and the impacts within the system. Each of these authors seeks to improve the operations of governmental and non-government organizations serving immigrants. One of the largest, and yet understudied, immigration operations in the United States is the immigration court system.

While research has examined the United States immigration court system, these works have mainly focused on using machine learning to examine the outcomes and bias related to asylum cases [287–289]. Similarly, Levesque [290] use analytics to examine the disparity between the duration and decision of cases within the United States immigration system. We contribute to this effort by modeling the United States immigration court system using discrete event simulation (DES). Rather than examine the equity of a past decision, we seek to use analytics to inform the decision-making process by evaluating current system throughput under varying conditions.

A few extant simulation models use DES to suggest improvements to immigration operations. Gandure and Mhlanga [291] use a DES model to examine the port-of-entry immigration operation in Botswana. Lim and Nor [292] apply DES to model a governmental office for immigration services (such as processing passports and visas) in Malaysia. While both works examine the operations of immigration services using DES, these works model fairly straightforward processes where an individual enters and exits within the same day. Our work differs as we model the complex journey of cases throughout the case duration in the immigration court system, which could be years and involve several hearings. To the best of our knowledge, this is the first work to model the United States Immigration court system.

4.2.1. Immigration Court System

Six different federal agencies handle immigrant-related affairs in some capacity [248]. One of the largest agencies is the Executive Office of Immigration Review (EOIR), a division of the Department of Justice (DOJ). The EOIR is tasked with processing all immigration cases involving formal legal hearings in the United States. The EOIR comprises two entities: i) The Office of the Chief Immigration Judge (OCIJ) and ii) The Board of Immigration Appeals (BIA). OCIJ handles the general immigration court proceedings. The BIA is separate from the OCIJ as this entity is specifically in place to review decisions that come as a result of an appeal filed for an OCIJ case; in a way, the BIA acts as the "checks and balances" of the immigration court system, and for this reason, the way that the OCIJ and BIA systems operate are distinctly different. The OCIJ is fashioned in what most would consider a typical civil legal setting, in which individuals make claims or defend themselves against a claim, and a judge decides how to adjudicate each case. The BIA does not function in similar fashion; thus, our work solely models the OCIJ court process. From this point, we will refer to the OCIJ simply as the immigration court [248, 293]. Many factors impact the operation and, therefore how cases flow through the system, and are described below.

4.2.2. Immigration Judges

The process and system structures of the immigration court differ from that of criminal and civil courts. While immigration cases must also abide by federal laws and can receive criminal convictions, unlike criminal and civil courts, the verdict of a case is determined entirely by the immigration judge presiding over each case rather than a jury of peers. Immigration judges primarily oversee cases pertaining to deportation, asylum, and other forms of relief available for persecuted individuals. Immigration judges do not have the same judicial independence and life tenure as other federal judges, and the requirements for becoming an immigration court judge are relatively minimal. Unique compared to other federal judges, immigration judges can be fired and hired more freely, creating a larger turnover in judge tenure [294, 295]. Given the more flexible tenure of immigration judges, over the last several years, additional judges have been hired in an effort to reduce the growing backlog [296]. At the end of the 2019 fiscal year, the EOIR had 442 active immigration judges across 68 different locations [297, 298]. The fluctuation in tenure and the number of active judges over time impact the modeling approach of judges, which is described further in Section 4.3.

4.2.3. Court Proceedings

In the immigration court context, a court proceeding is a formal process where the Department of Homeland Security (DHS) charges an individual with violating immigration law. Each immigration case that processes through the immigration court is comprised of at least one court proceeding; however, an individual may face multiple violations and have multiple proceedings over time. Each proceeding covers a different violation and is concluded when a verdict is reached; therefore, in our context, we model each proceeding independently. In other words, we model an individual's journey through the system on a "proceeding level" rather than on a "case level" (Figure 4.1). We consider a case to be completed when a proceeding reaches a verdict rather than when *all* the associated proceedings of an individual have reached a verdict and henceforth use "case" and "proceeding" interchangeably. Each proceeding is generally comprised of multiple appearances before a judge at what is known as a scheduled calendar hearing.

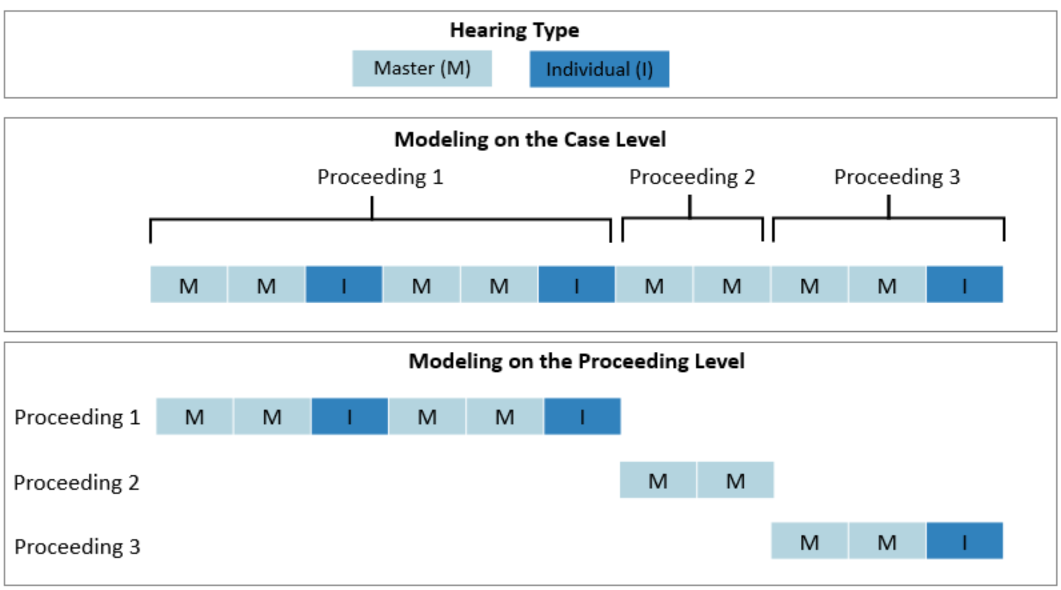

Figure 4.1: **Modeling Immigration Cases on "Case" verses "Proceeding" Levels.**

4.2.4. Scheduled Calendar Hearings

In the immigration court, a scheduled calendar hearing is when a case has a hearing before an immigration judge regarding its proceeding. The two main types of calendar hearings are Master and Individual hearings. In most circumstances, the first hearing is a Master calendar hearing, while the last is often an Individual hearing. Master calendar hearings are short preliminary hearings (typically 15 minutes) used to deal with the scheduling and procedural aspects of a case. Master calendar hearings prepare the defendant, attorney, and judge for the formal Individual hearing [263].

Individual calendar hearings are longer (usually four hours) and used to address the substance of charges against an individual. These hearings provide the opportunity for each individual (and their legal counsel) to plead their case by providing evidence and testimony. Individual calendar hearings typically conclude with a verdict from the immigration judge [263]. As mentioned previously, each proceeding comprises multiple calendar hearings, which we refer to as a case's *sequence*. The unique aspects of each case impact their sequence and contribute to a case's total time in the system, which we refer to as *sojourn time*. Figure 4.2 displays the intuitive increasing relationship between the sequence length and total sojourn time for completed cases in the New York City immigration court from 2010-2019. One aspect that significantly impacts the sequence and, therefore, sojourn time of a case is the case's priority status.

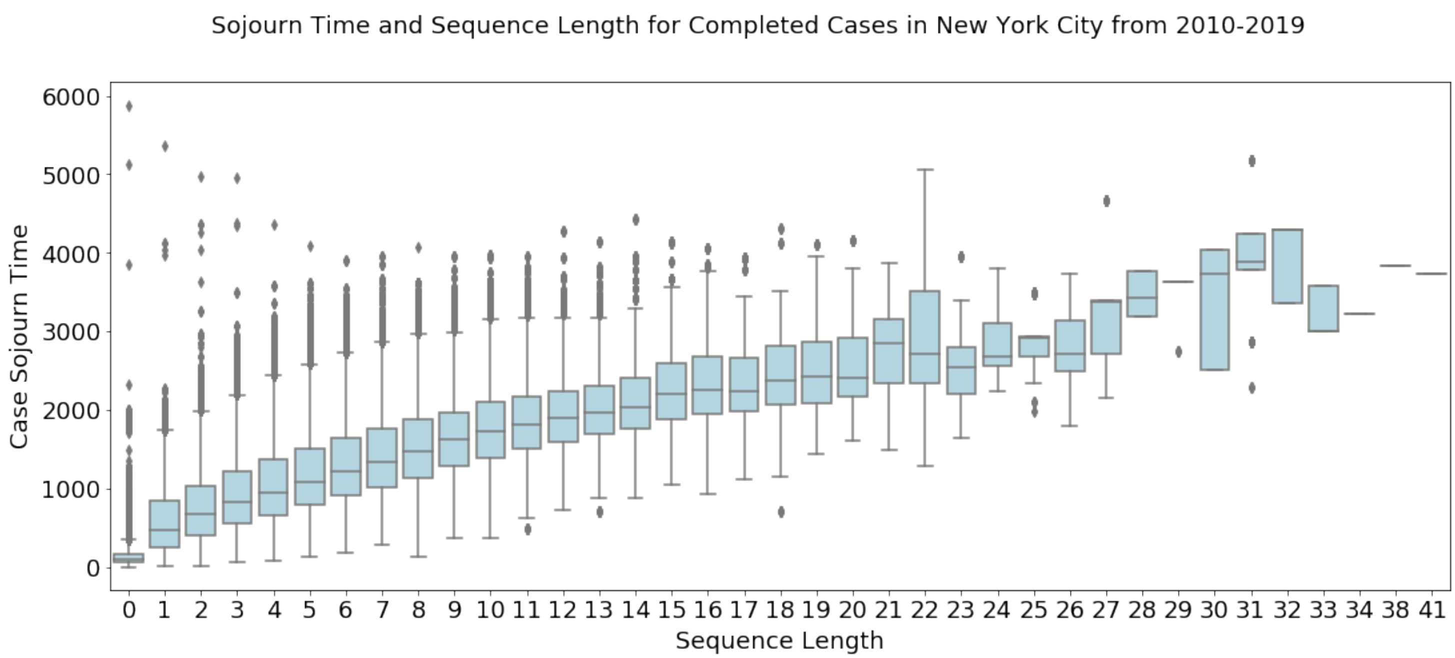

Figure 4.2: **Comparison of Sojourn Time and Sequence Length for Cases Completed in the New York City Immigration Court from 2010-2019.**

4.2.5. Case Priorities

In the immigration court context, there are varying levels of case priority. The policies determining which cases have priority over others change frequently [253,261]. Because of fluctuating policies, it can be difficult to attribute the priority status of most cases with one notable exception: asylum seekers. Asylum-seeking cases are typically given priority over other cases, a policy which has remained constant over time. In the United States, asylum is a form of protection that allows individuals with a "well-founded" fear of persecution in their home country to remain legally in the United States. Distinctly different from refugees, who gain legal status *prior* to entering the United States, asylum seekers are only eligible to apply for asylum once they are *physically present* in the United States [259]. In the EOIR immigration court, an asylum application most often occurs *after* an individual has already begun their proceeding and is a particular type of asylum known as defensive asylum. Defensive asylum applications can occur at any point in time within the first year of an individual entering the country. Policy states that asylum seekers are entitled to an expedited process [253, 299]. While not all eligible cases exercise this right, the expedition of asylum cases takes priority over other cases. Therefore in our context if a case has submitted an asylum application [1] at any point in time during their proceeding, we classify the case priority as "Asylum", and "Non-asylum" otherwise.

A case's priority status impacts how a case progresses through the court system, and also affects how quickly a case is heard. For example, in 2018 expedited asylum cases were expected to be adjudicated within 180 days and in 2020 detained cases were given priority and expected to be adjudicated within 60 days [261]. Given the limited resources of the immigration court, case prioritization can cause additional delays to nonpriority cases, contributing to the complexity of scheduling cases.

4.2.6. Case Scheduling

One of the most intricate modeling elements of the immigration court process is the scheduling of cases. While some aspects of the scheduling process are publicly documented [263], the guidelines have evolved over time and are often left to the discretion of judges and court staff. Upon arrival to the immigration court system, three main elements dictate how cases are scheduled: service demand, service capacity, and scheduling rules. The service demand is the number of new cases each day requesting service from the immigration court, which we consider to be the *arrival rate*. The service capacity is the number of cases each court can process on a given day, which we consider to be the *service rate*. The scheduling rules, or *queueing discipline*, determine the order in which cases should be served – the relative priority of cases. To understand and represent these complexities, we use historical data from the EOIR immigration court to inform the arrival rates, service rates, and queue discipline.

[1] We consider applications for "ASYL", "ASYW" and "WCAT" under the wider classification of asylum. We refer the reader to [300].

4.2.7. Immigration Court Data

Starting in 2016, anonymized data containing all immigration cases handled by the EOIR have been made publicly available through the Freedom of Information Act (FOIA) [276]. This data is the EOIR case dataset (CASE) and contains enriched information on varying aspects of each case across all court locations from 1950 to the present (see 3.6). While CASE data through February 2023 are available at the time of this writing, COVID-19 caused severe court closures and temporary changes to processes resulting in abnormal system behaviors. In addition, while all cases progress through the EOIR system as a whole, there exists variation in processes and behaviors across locations. These variations can be due to a number of reasons, such as the the Assistant Chief Immigration Judges (ACIJs) who oversee each court [259, 301], the court's staff and available resources.

Considering the variation across courts and the impacts of COVID-19, we focus the present analysis on modeling New York City (NYC) immigration court from 2010-2019. The rationale of selecting a single site are (i) to first establish a model which is stable, (ii) capable of handling the large complexity of the system, and (iii) to capture the magnitude of the backlog. By establishing a foundational model that is well-calibrated to empirical values, we can later expand our model to include additional cities. For this reason, we selected to model the NYC immigration court, which has one of the largest backlogs for any single location [2] and well established (stable).

Section 4.3

Model Overview

Our model has two types of entities: i) Asylum cases and ii) Non-asylum cases. Figure 4.3 depicts the process flow in the immigration court for these entities. There are two processes (queues): Master calendar hearing and Individual calendar hearing. Entities arrive in the system and are assigned three attributes: entity type, sequence, and judge. The assignment of these attributes follows the assumption that the empirical data are representative and are assigned using the empirical probability mass function (PMF). Next, each entity is assigned to the appropriate queue depending on their sequence and judge. An entity's position in the queue is determined using a "ticketing system", which we describe below. Each entity continues through the system until they reach the end of its sequence, at which point the case exits the system. We next describe the system parameters and inputs.

Preloads and Warm-up. To model the system state between 2010-2019, cases that were already in the system prior to 2010 need to be accounted for. We define cases that start in the immigration court system prior to our simulated period as *pre-loaded* entities, and those entering during our simulation period are defined as *simulated*

[2] Note there are three immigration court locations within New York City, we are modeling the main court location denoted in the data by the code "NYC".

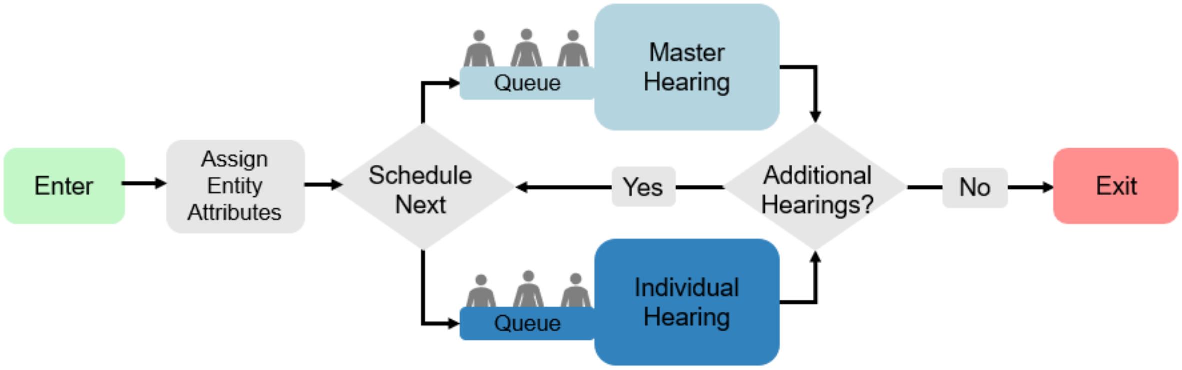

Figure 4.3: **United States Immigration Court Process Flow.**

entities. Modeling pre-loaded entities at the start of our simulation ensures that the existing system demand is captured. Because pre-loaded cases have already begun their proceeding, including pre-loaded cases adds bias when evaluating sojourn time. A warm-up period can minimize this bias by clearing out pre-loaded entities, leaving only simulated entities for evaluation. Considering the years 2000 through 2009, a warm-up period of five years (starting in 2004) minimized the total number of pre-loaded entities in the system at the start of 2010.

4.3.1. Sequence and Judge Assignment

Upon entry to the simulated system, each entity is assigned three attributes: i) priority status (Asylum or Non-asylum), ii) sequence, and iii) judge. Additionally, for each hearing (step) in a sequence, a binary attribute associated with court-caused delays is assigned. For example, a case may have a hearing scheduled; however, for reasons beyond the entity's control, the case is delayed and rescheduled. The hearing is then considered to be delayed – the delay time is the time between the day of notice for the delay and the rescheduled day. The assignment of each of these attributes follows the assumption that the sequences of cases and the associated characteristics (priority status and judge) are representative of the empirical data. For each new entity arriving in the system, the arrival year is used to generate a set of observed sequences from the empirical data, and from the selected subset, using the PMF a sequence is assigned. Subsequently, the set of all observed case profiles (priority status and judge) associated with the given sequence is uniformly selected from and designated to the entity. We consider sequences to be i) complete or ii) partial. A complete sequence refers to instances where an entity's complete sequence is known prior to exiting the system. In other words, complete sequences are attributed to cases that exited the system by the end of 2019. Partial sequences represent instances where an entity has not yet exited the system, and their sequence is not (yet) fully known. Sequences are assigned based on the arrival year of each entity, and therefore, cases with completed sequences decrease as the year increases (see Appendix B, Figure 17). This method of entity attribute assignment effectively preserves the behaviors and characteristics of the empirical data.

Arrival Rates. Arrival rates reflect the rate of new cases entering the immigration court system. We use empirical data to calculate the number of new cases arriving daily for the entire NYC court system. For each year, we generate a PMF from daily arrivals in the empirical data and probabilistically assign the arrival rates for each simulated day. Figure 4.4 shows the average number of arrivals for 2004-2019 (red horizontal line depicts the average across all years).

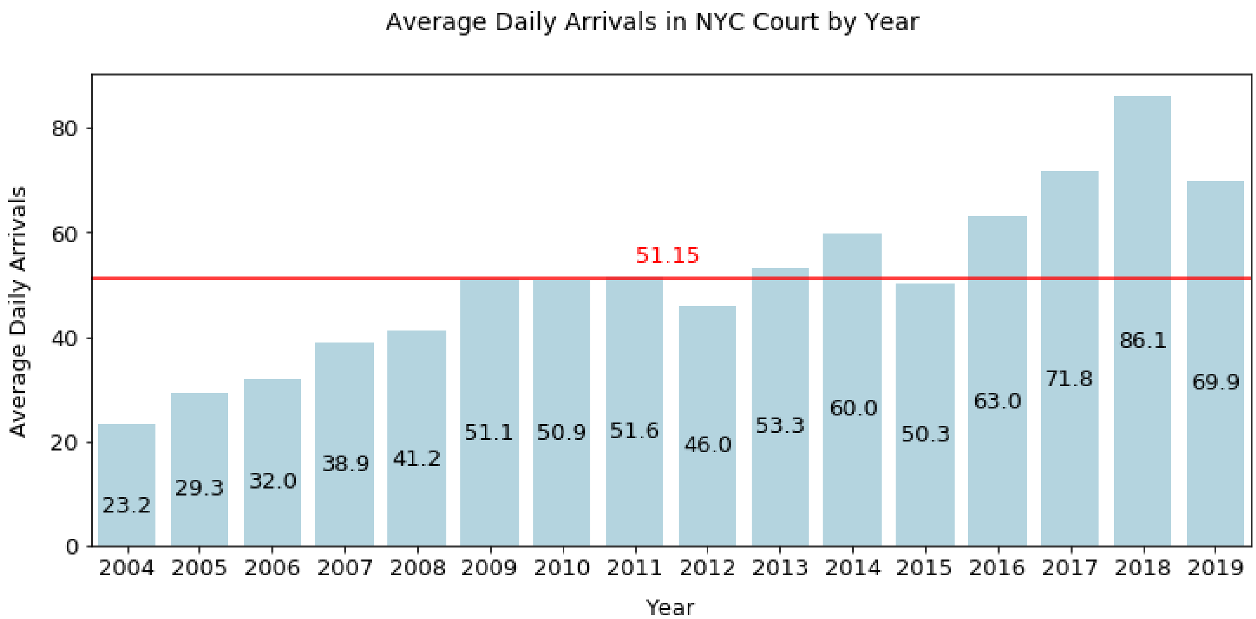

Figure 4.4: **Average Arrival Rates in New York City from 2004-2019.**

Judge Modeling and Service Rates. Due to the complexity and instability of judge availability and tenure over time, modeling individual judges is prohibitive and therefore, two methods of generalizing judges are considered (Figure 4.5). Although 139 unique judges are observed from 2004-2019, on average, only 23 judges are available on any given day. What we term as Model A represents this average availability through the uniform assignment of judges into 23 classes. What we term as Model B groups judges by average service rates into four classes (quartiles). Daily service rates are probabilistically assigned using a PMF stratified by year and hearing type. Figure 4.7 shows the average service rates for each queue for the years 2004-2019.

4.3.2. Scheduling Rules

Queue discipline, also known as a scheduling policy, refers to the order in which cases are served. In the immigration court context, the queue discipline is best characterized by a variation of the *Earliest Due Date* (EDD) scheme;

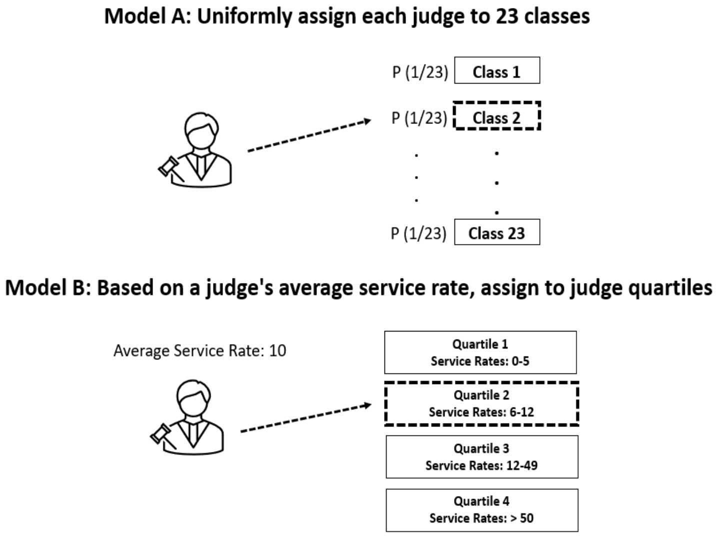

Figure 4.5: **Two Approaches for Modeling Judges.**

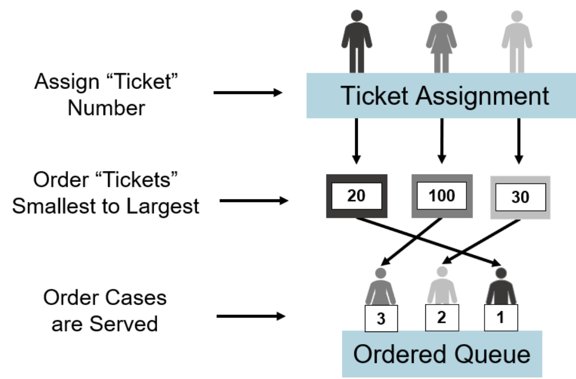

Figure 4.6: **Ticketing Approach to Assigning Cases to Queues.**

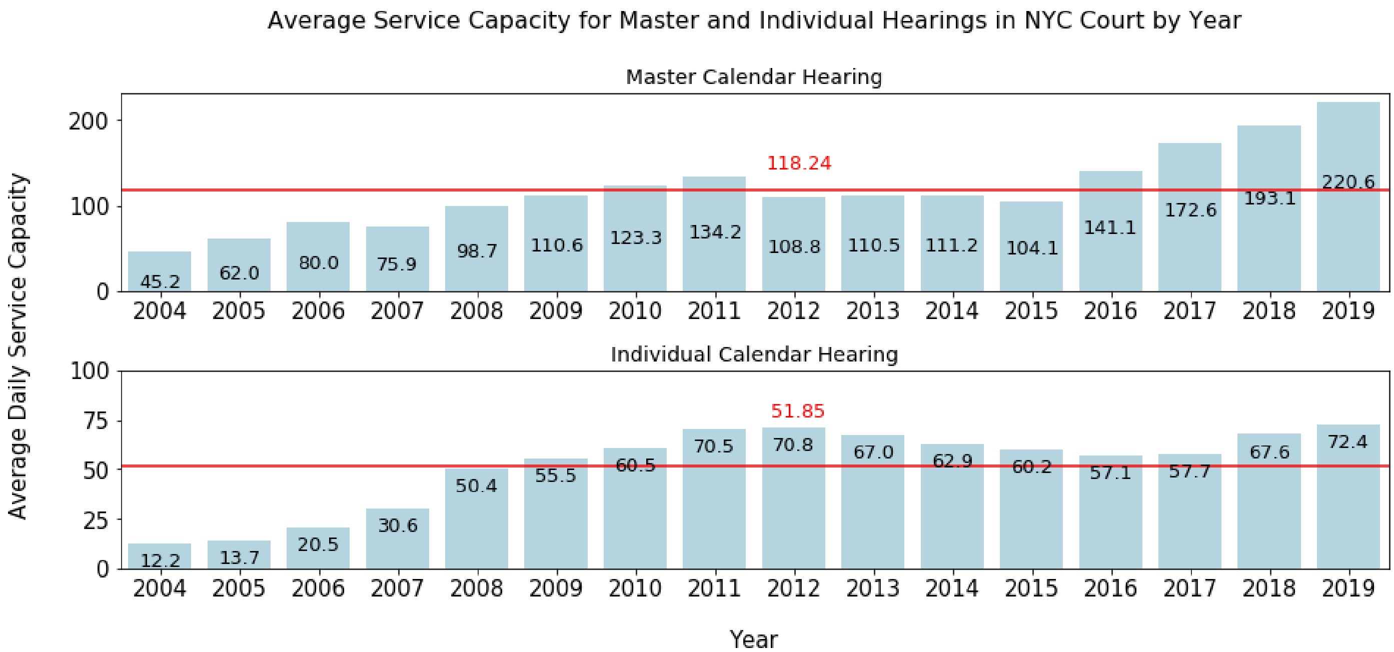

Figure 4.7: **Average Service Rates in New York City from 2004-2019.**

where cases with an earlier hearing date are served first. Each hearing date is scheduled in advance and determined by i) a set of scheduling rules and ii) on a first-come, first-served basis. Traditionally the queue discipline is clearly defined and stays constant throughout the simulated time. However, the scheduling rules of the immigration court are not well documented and frequently change – a situation difficult to represent as a static rule. To overcome this difficulty, we propose a "ticketing" approach to infer the order of each case. Utilizing the empirical wait times of cases, we assign a "ticket" to each entity as they arrive at a queue. The value of the assigned "ticket" is then used to reorder the queue in ascending order. It is important to note that the assigned "ticket" value does not represent how long a case will wait before being served; rather, it represents the relative priority of cases in the queue. Judges then serve cases with the lowest ticket number first. Figure 4.6 depicts this implementation. The “ticket” values are assigned using the PMF of empirical wait times, stratified on the entity type (Asylum or Non-asylum), hearing type, and if the current hearing will experience a court delay (as determined by the pre-assigned sequence).

Section 4.4

Baseline Model and Validation

Due to long run times, we ran 10 replications for each model with a five-year warm-up period, followed by an additional 10 years for analysis, simulating nearly one million hearings over 15 years. All experiments were run using SimPy 3.0.11 [302] and Python 3.7.3, under Linux Ubuntu 20.04.4 on Worcester Polytechnic Institute’s high-performance computing research cluster. The runtime (wall clock time) for each replication ranged from 1 to 2 days for the warmup period and 4 to 6 days for the simulated time.

We consider three key performance indicators (KPIs) to validate our baseline model: (1) the average time a case waits from entering the Master hearing queue to being served (Master hearing wait time), (2) the average time a case waits from entering the Individual hearing queue to being served (Individual hearing wait time), and (3) the total time a case spends in the system before their case is completed (sojourn time). During the simulated period of 2010-2019, Model A served on average 688,751 hearings, and 150,439 cases exited the system. Model B served on average 735,846 hearings and 160,356 cases exited the system. Figure 4.8 depicts the empirical number of cases that enter and *exit* between 2004-2019 (black line) with our simulated Models A (blue line) and B (orange line) across 10 replications. Visually, both models follow the same general increasing trend as the empirical data with a notably higher rate of growth for Model A. Model A served fewer hearings and had fewer cases exiting the system than Model B, leading to a faster-growing backlog. The more extensive backlog associated with Model A likely results from how capacity is assigned to each of the 23 judge classes. Considering judges are randomly assigned to one of the 23 classes, the variability of values in the PMF used to assign capacity (as noted in Section 4.3.1) is larger and led to an increased chance of either over or under-assigning capacity for each of the judge classes. Therefore, assigning judges using quartiles as done in Model B resulted in a tighter bound for the total number of hearings and cases served.

Tables 4.1 and 4.2 provide the absolute mean percentage error (AMPE) of each simulated model for case wait time and sojourn time, respectively. Each table reports the mean and maximum AMPE and 95% confidence interval across 10 replications for each case type. The *mean* wait time represents the average time each case waited at the Master and Individual hearings during their time in the system and is bounded by 11.5% and 7.2 % for Model A and B, respectively. Model A reports a higher bound for the Individual hearing with a value of 18.6% for Non-asylum cases and 21.8% for Asylum. Model B has a relatively low AMPE for most of the KPIs, however, has a high value reported for Asylum cases at the Individual hearing (29.4%).

While the responsibilities for cases during the Master calendar hearing tend to be more homogeneous such as filing paperwork, accepting pleas, and scheduling future hearings, the Individual hearing addresses the unique circumstances of each case in depth, such as providing testimony and producing evidence to support the defendants claims (see Section 3.2). Further, Asylum cases involving different groups may require different preparation times and hearing lengths, such as those involving family units or unaccompanied minors, and lead to larger variations in Asylum wait times at the Individual hearing. The empirical wait times for Asylum cases at the Individual hearing substantiate this observation with having the largest interquartile range (Appendix B, Figure 18). Assigning wait times for each case based on their individual case characteristics would improve the assignment of wait times for Asylum cases at the Individual hearing and is discussed further in Section 4.6.

The *maximum* wait time for all hearings is bounded by 36.3%. The larger AMPE for *maximum* wait times can be attributed to the maximum wait time in the empirical data being only six months less (14.5 years) than the total 15-year simulation period (including warm-up) as well as the previously mentioned unique circumstances experienced by Asylum cases at Individual hearings. Considering our simulation total run length is only 15 years, we simply may not have observed this large wait time yet. Finally, both the *mean* and *maximum* sojourn times (Table 4.2) are bounded by 14.1%. Overall, Model B performs better on both wait time and sojourn time when compared to Model A, and therefore, Model B is preferred in our context.

Table 4.1: **Absolute Mean Percentage Error for Case Mean and Maximum Wait Times across 10 Replications.**

Hearing Type	**Case Type**	**Model A**		**Model B**	
		Mean	Maximum	Mean	Maximum
Master	**All**	10.14% ±5.22%	4.11% ±3.01%	6.89% ±5.36%	4.74 % ±4.41%
	Non-asylum	11.5% ±5.83%	5.10% ±3.73%	6.50 % ±5.30%	4.26% ±3.95%
	Asylum	9.23% ±4.96%	3.91% ±3.14%	7.15% ±5.43%	5.165% ±4.71%
Individual	**All**	15.94% ±2.31%	23.63% ±2.01%	24.08% ±2.33%	31.35% ±1.94%
	Non-asylum	18.63% ±3.13%	7.33% ±2.67%	7.16% ±3.21%	3.48% ±2.45%
	Asylum	21.79% ±2.18%	29.13% ±1.91%	29.37% ±2.21%	36.30 ±1.84%

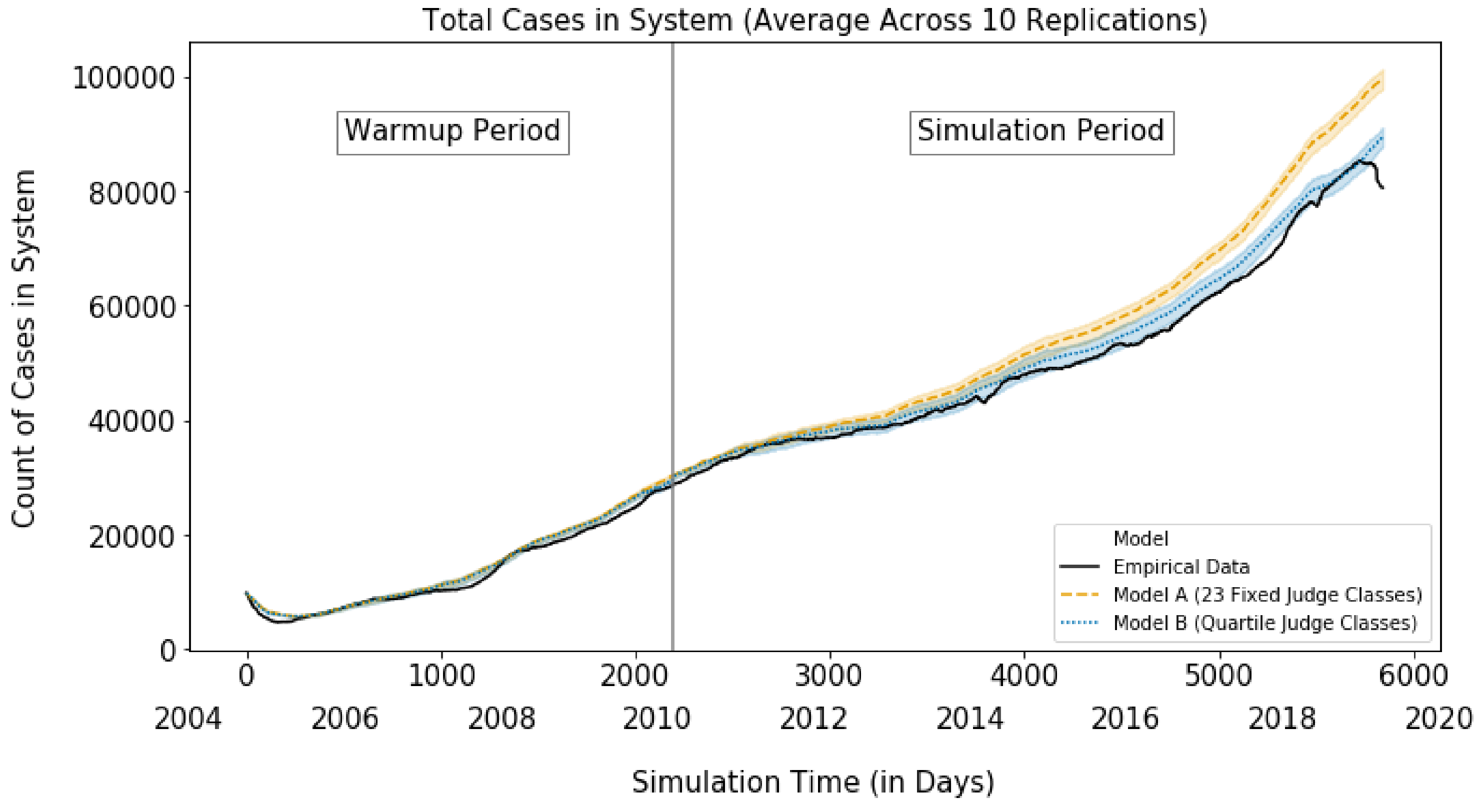

Figure 4.8: **Model A and B Simulation Output across 10 Replications.**

Table 4.2: **Absolute Mean Percentage Error across Models for Case Mean and Maximum Sojourn Times for 10 Replications.**

Case Type	**Model A**		**Model B**	
	Mean	Maximum	Mean	Maximum
All	4.73% ±2.54%	7.37% ±0.96%	3.11% ±2.98%	4.32% ±1.52%
Non-asylum	6.58% ±2.93%	2.33% ±2.31%	10.58% ±5.08%	1.46% ±1.13%
Asylum	10.22% ±1.82%	14.11% ±1.02%	9.41% ±2.72%	10.87% ±1.61%

Section 4.5

Sensitivity Analysis

In our baseline models, we now explore the sensitivity and impact of varying two model inputs: arrival rates (AR) and service rates (SR). We consider eight experiments depicted in Table 4.3. The variation for ARs and SRs was determined using the average yearly change observed in the empirical data. The average change for ARs was determined to be approximately ± 4%. The average rate of change for SRs for the Master and Individual hearings was determined to be ± 8% and ± 4%, respectively. The baseline rates are the same AR, and SRs generated for the baseline models as described in Section 4.3. We ran 10 replications of each sensitivity experiment for our two models, totaling 160 experiments.

Table 4.3: **Arrival Rates and Service Rates for the Eight Sensitivity Experiments.**

Experiment Description	Experiment Number	AR	AR Direction	SR M	SR I	SR Direction
Vary Arrival Rates (AR) Only	1	4%	+	Baseline Rates		
	2		-			
Vary Service Rates (SR) Only	3	Baseline Rates		8%	4%	+
	4					-
Same Direction	5	4%	+	8%	4%	+
	6		-			-
AR +, SR -	7	4%	+	8%	4%	-
AR -, SR +	8	4%	-	8%	4%	+

4.5.1. Results

Figures 4.9 - 4.12 compare the baseline total number of cases in the system from 2004-2019 for each of the eight experiments. The left column of plots shows the results for Model A, and the right column shows the results for Model B. Figure 4.9 depicts the experiments for varying arrival rates, where the green (dotted) line represents increasing the arrival rates by 4% and the red (dashed) line shows decreasing by 4%. As one may expect, when arrival rates increase (keeping service rates constant) the total number of cases in the system also increases. Conversely, when the arrivals decrease the total cases in the system also decrease. Considering the varying service rates experiments (Figure 4.10), the opposite expected behavior is observed; an increase in service rates reduces the number of cases in the system, and a decrease in service rates increases the total cases in the system. Examining the impacts of Experiments 5 and 6 (Figure 4.11), we observe that by increasing (or decreasing) the arrival and service rates in the same direction, the changes to ARs and SRs effectively cancel one another. Figure 4.12 shows how moving the rates in the opposite direction generates large deviations from the baseline models.

Figures 4.13 and 4.14 show the impact these variations in ARs and SRs have on average case wait times (shown across all hearing types) and case sojourn times. Comparing these two figures reveals that variations in service rates have a larger magnitude of impact on both average wait and sojourn times of cases. For example, the difference in sojourn time between the baseline model and Experiments 2 and 3 shows a smaller reduction in mean sojourn times when decreasing arrival rates than when increasing service rates. This suggests extra capacity in the system has a larger impact on KPIs, in comparison to additional demand coming into the system. Experiment 8 produces the largest reduction in both wait and sojourn times, whereas Experiment 7 produces the largest increase in wait and sojourn times.

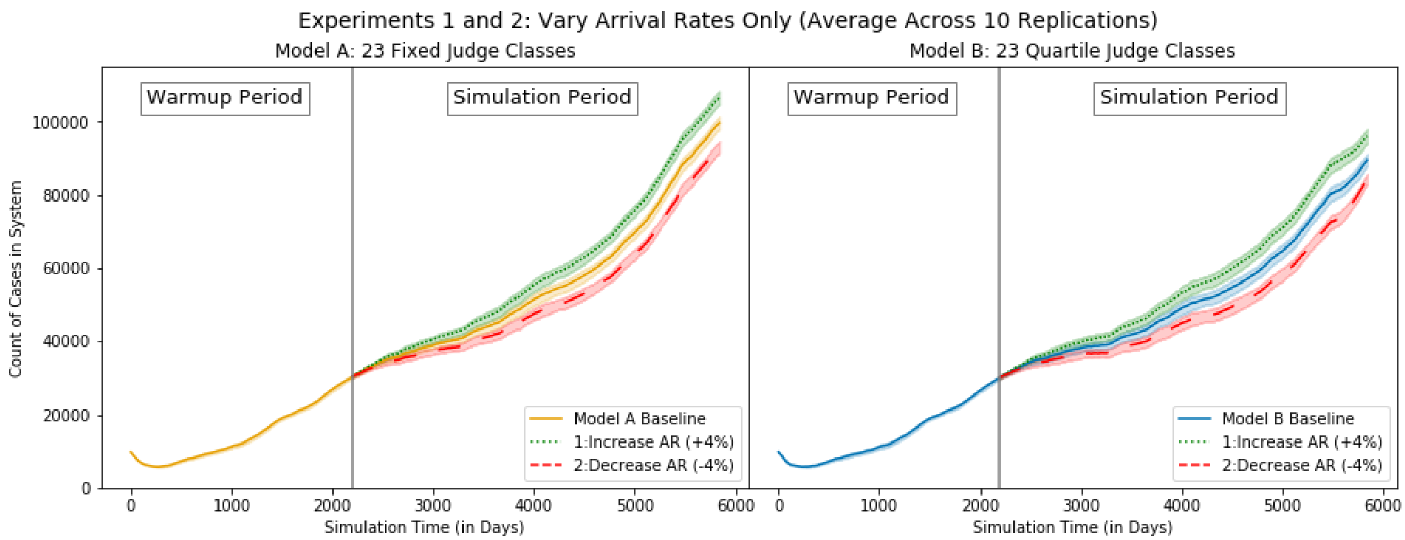

Figure 4.9: **The Total Number in System across 10 Replications for Varying Arrival Rates (Experiments 1 and 2).**

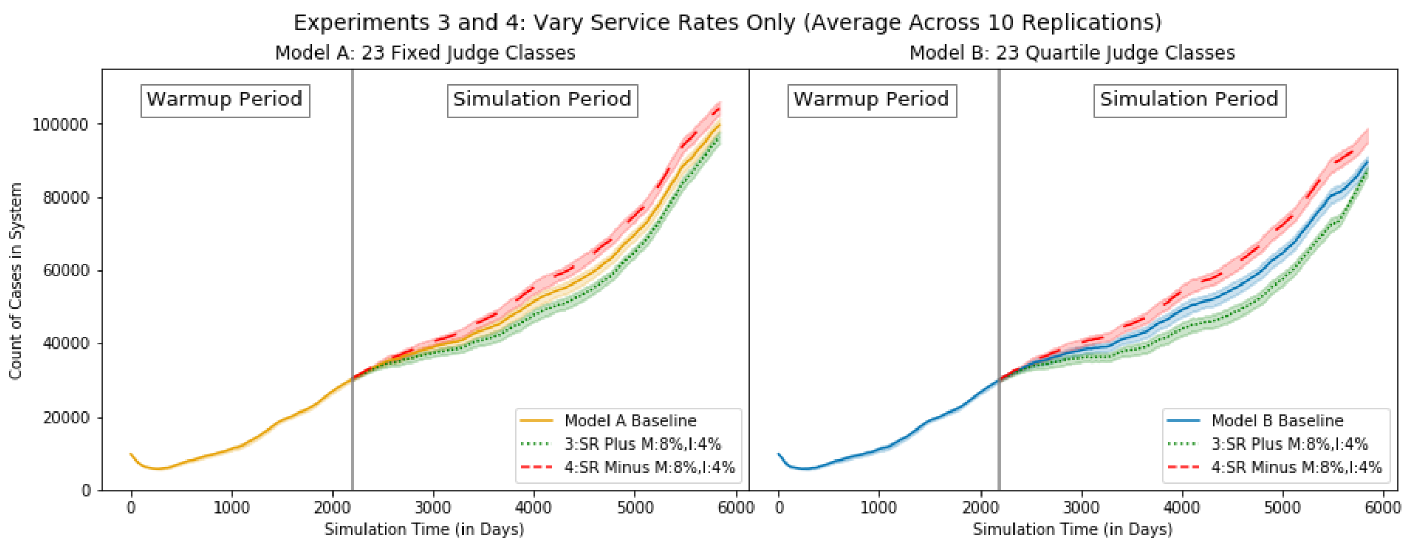

Figure 4.10: **The Total Number in System across 10 Replications for Varying Service Rates (Experiments 3 and 4).**

Although our results are intuitive, these experiments demonstrate two important aspects of our model: i) our model behaves in the expected manner (that is, when arrivals increase, the total number in the system also increases) and ii) our model can be used to assess real-life changes to arrival and service rates. For example, if there is an expected increase in arrivals, decision makers can use the present model to determine at what rate would the immigration court need to operate to reduce or keep constant the size of the backlog. Alternatively, decision makers could evaluate the level of additional resources (such as increasing service rates) required to clear the entire backlog while considering the impact on wait and sojourn times.

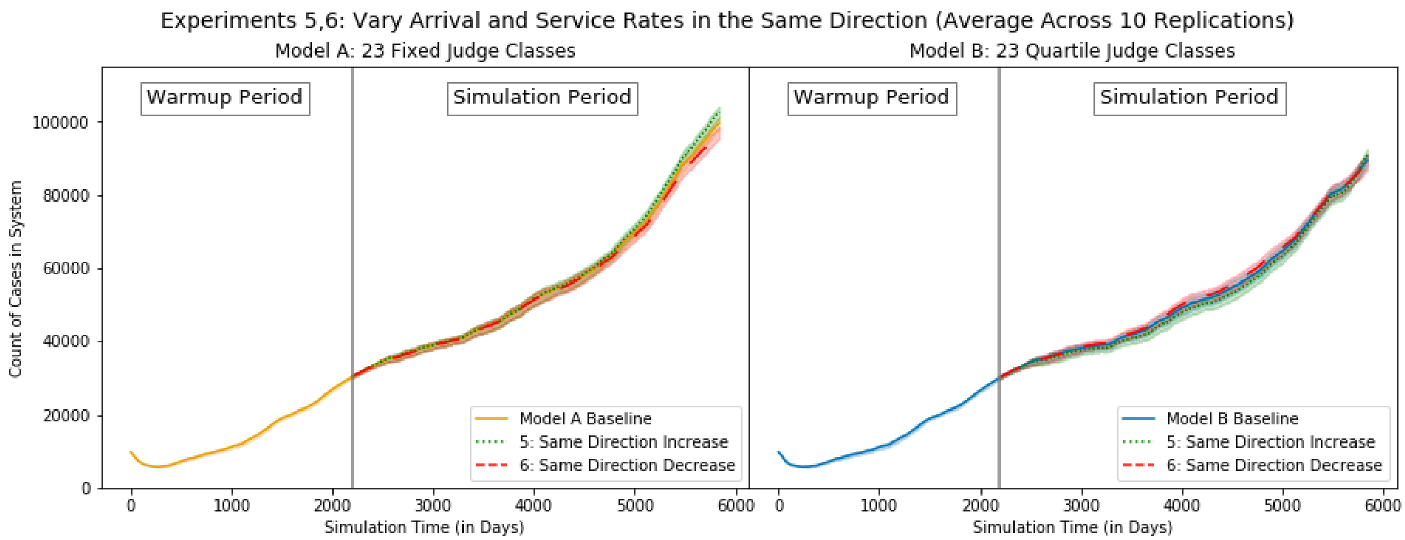

Figure 4.11: **The Total Number in System across 10 Replications for Varying Arrival and Service Rates in the Same Direction (Experiments 5 and 6).**

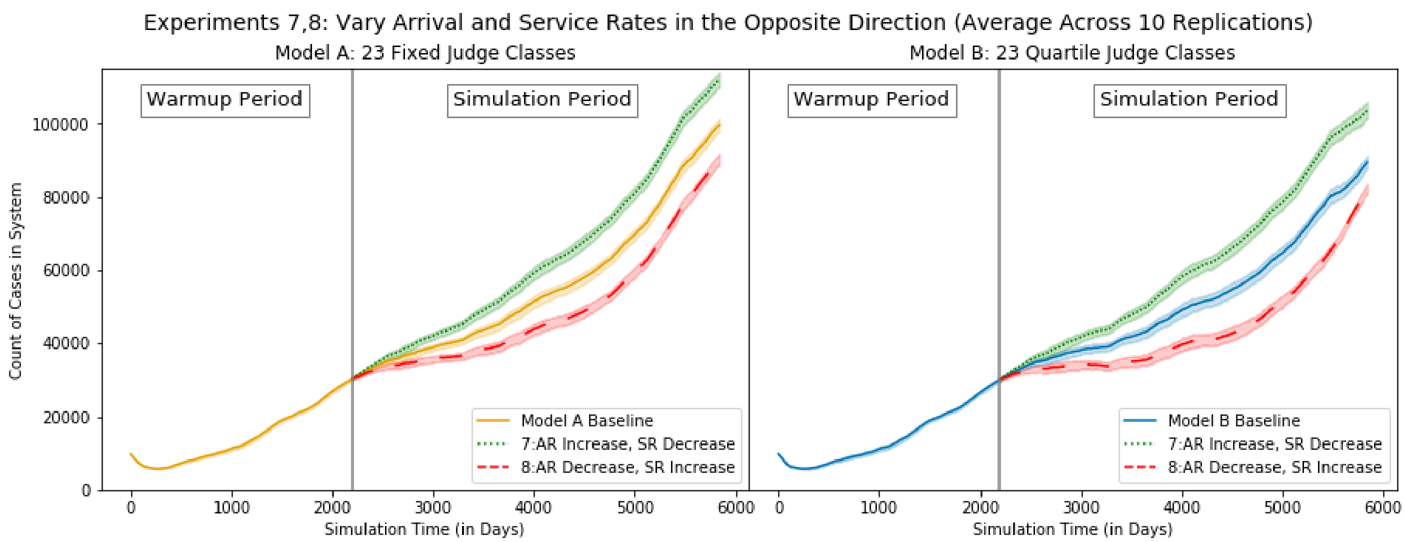

Figure 4.12: **The Total Number in System across 10 Replications for Varying Arrival and Service Rates in the Opposite Direction (Experiments 7 and 8).**

Section 4.6

Conclusion and Future Work

This chapter presents a discrete event simulation model of the United States Immigration court system. Using 15 years of historical data (a five-year warm-up period, followed by an additional 10 years), we simulate daily arrival rates, service capacities, and case sequences, capturing the variation of this complex system. Simulating nearly one million hearings within the New York City court, we establish and validate a baseline model. Building off the baseline model we examine the system sensitivity to changes in arrival rates and service capacities, demonstrating

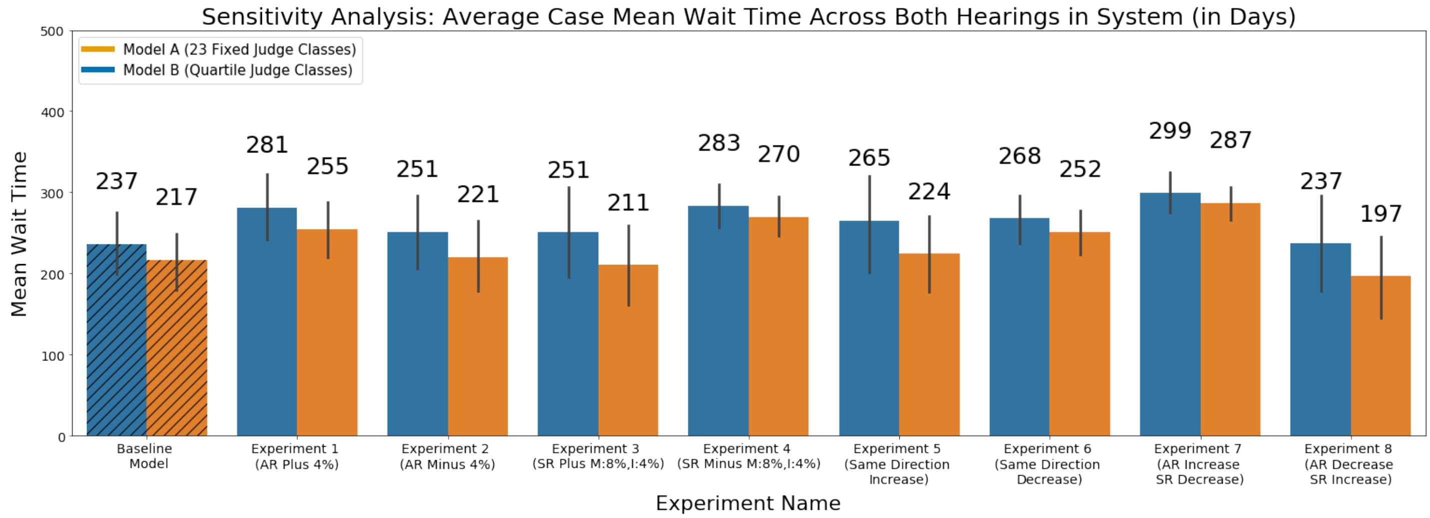

Figure 4.13: **Average Wait Time across 10 Replications and 8 Sensitivity Experiments.**

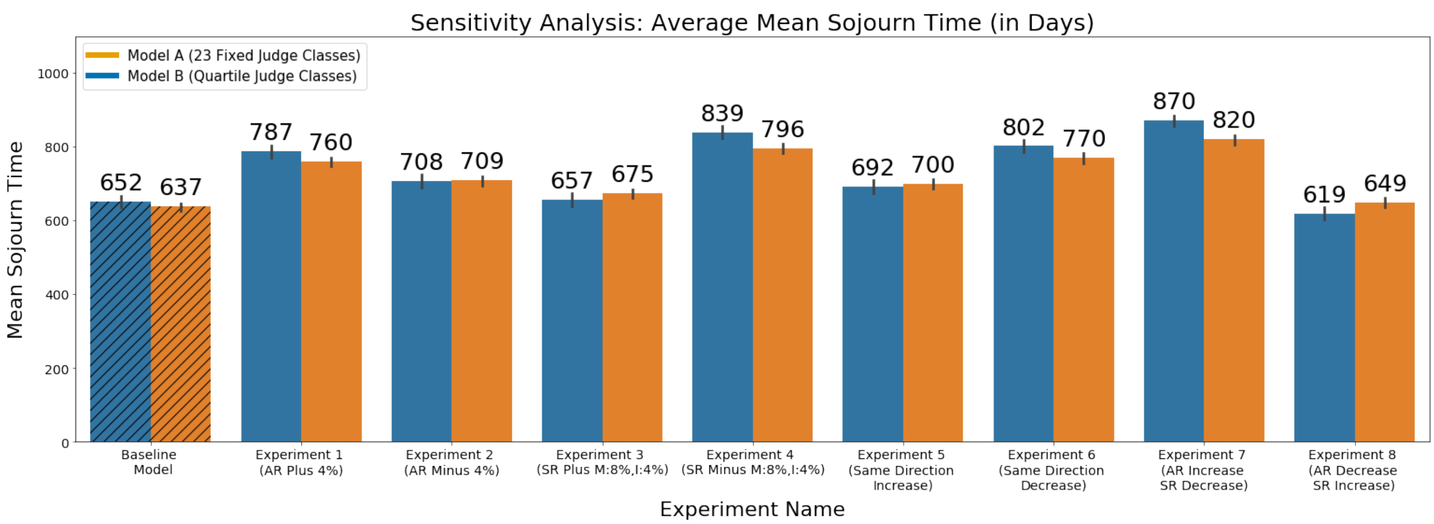

Figure 4.14: **Average Sojourn Time across 10 Replications and 8 Sensitivity Experiments.**

the effectiveness of our model in evaluating system KPIs.

Through the development of our model, we provide an analytical representation of the United States immigration court system, examining the current system throughput and evaluating the impact the growing backlog has on different case types. Modeling a large and complex system such as an immigration court system is not without limitations. Our current model generalizes Asylum and Non-asylum cases, assuming homogeneity within case type. As shown in Table 4.1, this assumption contributed to a large mean percentage error for Asylum wait times at the Individual hearing. In addition, while our model represents the general state of the system well, the present analysis is limited to examining only the impacts that variations in arrival and service rates have on the behavior of the system.

Future work can address these limitations. For example, through the use of a machine learning model, our "ticketing" system can be improved to consider case characteristics and improve the representation of heterogeneity within Asylum and Non-asylum case types. Additionally, future work can expand our model to handle more complex changes to the system that can evaluate, inform and improve immigration court policies. Our model is unique, as we can examine the impact policy has on reducing the immigration court backlog, as well as assess the equity of such a policy by examining the effect each policy has on individual cases. For example, while a policy prioritizing Asylum cases may effectively reduce the immigration court backlog if the same policy triples the maximum wait time for Non-asylum cases, decision makers would benefit from knowing this. Chapter 5 expands the presented model, incorporating three policies informed by our analysis in this chapter and supported by domain experts, seeking to demonstrate the impact of this model further.

Chapter 5

Chapter 5: Assessing Policies to Improve the United States Immigration Court Operations Using Discrete Event Simulation

Section 5.1 Introduction

The work presented in Chapter 4 highlights the complexity, interdependency, and sensitivity to changes in arrival and service rates in the New York City (NYC) immigration court from 2010-2019. Informed by these insights, in this chapter we extend the current simulation model to enable an in-depth and robust exploration of how policy changes can impact and improve outcomes. We investigate three policies informed by our analysis and supported by domain experts and evaluate the impact of each on improving three key performance indicators (KPIs)– sojourn times, wait times, and queue lengths. In particular, we examine the policies of: i) vary the number of new judges, ii) dedicated dockets for asylum cases and iii) make-up capacity for delayed cases. The first policy evaluates the "value-added" of hiring or adding new judges to serve cases in the NYC immigration court system. The second policy implements two new dockets (one for each hearing type) designed to serve asylum cases more expeditiously. We assess the impact these dedicated dockets (queues) have on the equity of service times within and between case types. The final policy deploys a novel make-up capacity scheme designed to reduce the time added for cases that experience court-caused delays.

The implementation and evaluation of these three policies provide case-level insights for decision makers and illustrate the power of the data-driven Discrete Event Simulation modeling. This analysis lays the foundations for an adaptable and interactive decision-making tool. The deployment of such a tool can provide stakeholders with a tangible way to seek more efficient and equitable solutions to the growing backlog and better serve those in the immigration court system.

The chapter proceeds as follows. Section 5.2 introduces, provides context, and motivates the three proposed policies. Sections 5.3, 5.4, and 5.5 provide the details on the implementation of each policy and the ensuing results and evaluation. Finally, Section 5.6 concludes the work with a discussion and directions for future work.

Section 5.2

Background and Policy Context

5.2.1. Policy 1: Varying the Number of New Judges

To expedite the reduction of the growing immigration court backlog, the Executive Office of Immigration Review (EOIR) has sought to hire an increasing number of judges over the last several years. While in 2010 the EOIR had approximately 235 immigration judges across 59 locations [296], by 2019 there were 442 immigration judges with 92 new judges hired in that fiscal year (FY) alone[1] [297], [303]. A closer look at the data presented in Chapter 4 shows how the increased hiring of EOIR judges is reflected within the NYC court context. Figure 5.1 shows the average unique number of active judges for FY 2004 to 2019. Figure 5.2 depicts the total count of new judges observed each FY from 2004 to 2019. These figures illustrate an increase in both the hiring of new judges and the active number of judges across the 15-year time horizon.

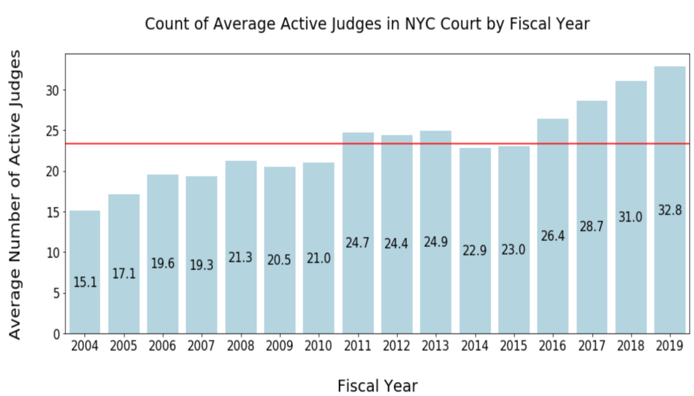

Figure 5.1: **Average Number of Daily Active Judges in the New York City Court System for Fiscal Years 2004-2019. The red line depicts the average across the time horizon.**

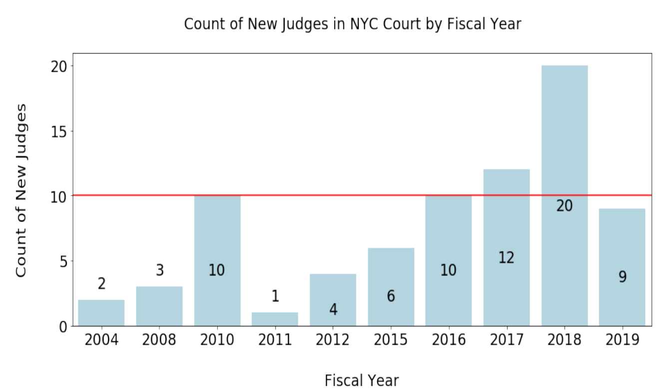

Figure 5.2: **Number of New Active Judges in the New York City Court System for Fiscal Years 2004-2019. The red line depicts the average across the time horizon.**

As shown in Chapter 4, one can intuitively see that by increasing the number of judges, the total number of cases served would also increase, producing similar outcomes to those presented in section 4.5 (Figures 4.10,

[1] A Fiscal Year (FY) in the United States Federal Government lasts from October 1st - September 30th.

4.12). A 2018 study by the Bipartisan Policy Center confirms this conjuncture and evaluates three proposed hiring schemes and their association backlog reductions [304]. While these projections demonstrate the potential benefits of hiring more judges, the modeling effort failed to incorporate the vast complexity of the court process and idiosyncrasies of case trajectories limiting subsequent results (Chapter 3, Figure 3.9). The aforementioned model considers the aggregate values of cases served per year, per judge, and the number of arrivals per year, examining only the impacts of overall queue lengths. Importantly our model presented in Chapter 4 considers the variation in case types and sequences across judges and years. Thus modeling a policy of increasing judges using the present model in Chapter 4 enables a more comprehensive look at the effects on the backlog, as well as impacts on case wait and sojourn times.

The operational and ethical impacts of hiring more judges warrant careful consideration. Operationally, a limited number of courtrooms, staff, and finances are available to support hiring new judges. While increasing the number of judges can theoretically serve to reduce the backlog, in practice, the court must consider the resources required to support expansion. Ethically, concerns exist that hiring large quantities of judges to expedite reductions to the backlog may process cases too quickly and impede an individual's right to due process [305]. Processing cases too quickly may reduce an individual's ability to obtain legal counsel and adequately prepare for their hearing [306]. Additional concerns arise from the fact that requirements for becoming an immigration court judge are relatively few, which can result in hiring judges without prior immigration law experience [305]. While the EOIR currently provides six weeks of training [307], there are concerns that given the complexity of immigration law, such brevity can not fully train judges to understand and implement to the level of adequacy for due process [305]. Considering these concerns, it is essential for policy and decision makers to have a deeper understanding of not only how increasing the number of judges can expedite reductions to the backlog (queue lengths) but also how hiring more judges impacts the cases that process through the system and is explored in Section 5.3.

5.2.2. Policy 2: Dedicated Dockets

Dedicated dockets act as new, restricted queues serving only certain case types. Over the past several years, dedicated dockets (informally known as "Rocket Dockets") have been proposed and implemented within the EOIR as a method to expeditiously and fairly process cases [308]. Dedicated dockets have been deployed to serve a variety of case characteristics such as unaccompanied minors [309], adults with children (AWC) [310], and most recently, asylum seekers; specifically, families arriving along the Southwest border of the United States [306].

Concerns exist regarding the effectiveness of expediting cases through dedicated dockets and the immigration courts' ability to adhere to the guidelines and objectives of such. For example, at the end of May 2021, the Biden administration announced the deployment of the dedicated dockets program meant to serve families arriving at the

Southwest border of the United States. The associated guidelines and objectives were to reduce the total sojourn time of cases to less than 300 days (from the time of the case's initial master calendar hearing to the verdict) while maintaining fairness. Analysis conducted by Transactional Records Access Clearinghouse (TRAC) [280] found that as of December 2022, of the over 110,000 cases assigned to the docket, nearly 40,000 reached completion. The analysis found that while 83% of cases were closed in less than 300 days, the fairness and impact on case outcomes due to the expedited process remain a concern [306]. For example, the analysis found that only 7% of the cases on the dedicated docket were granted asylum, and only 34% of the cases found legal representation. TRAC further postulates that a reduction in the immigration court backlog (in particular asylum cases) can be fast or fair, but not both. Many other immigration advocates share the same concerns regarding fairness in the use of dedicated dockets to serve cases expeditiously [309], [311], [312]. Thus, evaluating dedicated dockets in terms of fairness is a critical dimension to consider. While the due process and fairness of a case in terms of case outcomes (whether relief was granted) are challenging to measure and assess, the trade-off of cases expedited through a dedicated docket compared to the general docket can proffer insights into the equity of such a policy. For these reasons, we deploy dedicated dockets for asylum cases [2] and explore the impact that shifting capacity resources has on system KPIs in Section 5.4.

5.2.3. Policy 3: Make-up Capacity

The extreme backlog and prolonged wait times in the immigration court are a function of many factors, including a large influx of immigrants coming from the Southwest border of the United States, an increase in interior deportations (resulting in large quantities of cases entering the system), the deployment of the Migration Protection Protocol (more informally known as "Remain in Mexico") and more recently the COVID-19 shutdown [313], [314]. In addition to these factors, hearing delays further impede the court's ability to adjudicate cases in a timely manner. Approximately 30% of all hearings in the NYC immigration court from 2004-2019 experienced some type of delay [3], and nearly 20% of all hearings experienced a delay due to the Department of Justice (DOJ) or the Department of Homeland Security (DHS). These court delays can occur for various reasons, such as prioritization for other cases and over-scheduling. Figure 5.3 shows the percentage of cases delayed due to the applicant, DOJ, and DHS out of all case delays. More than half of the delays that occurred were associated with non-applicant delays. Table 5.1 provided the mean and maximum wait time "added" to cases due to each delay type.

[2] While the policy as of May 2021 surrounding dedicated dockets prioritizes family units seeking asylum at the Southwest border of the United States, we broaden the use of these dockets to all cases who have submitted an asylum application at some point in time during their proceeding

[3] We assign delays using the adjudication codes associated with each hearing

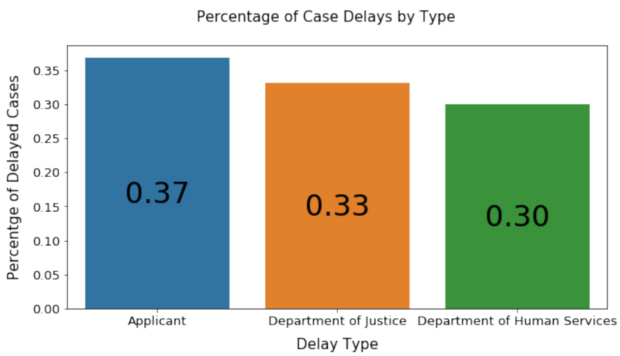

Figure 5.3: **Percentage of Case Delays by Type in the NYC Court System for FY 2004-2019.**

Table 5.1: **Mean and Maximum Wait Times Added (in Days) per Delay Type.**

Delay Type	Wait Time Added	
	Mean	**Maximum**
Applicant	194.5	4,598.0
Department of Justice	253.0	2,037.0
Department of Human Services	223.7	3,856.0

Seeking to reduce the additional wait time resulting from DOJ or DHS-initiated hearing delays, we implement a novel scheduling scheme that reserves "make-up" capacity semi-regularly for hearings impacted by court-caused delays. This policy investigates the influence of reserving capacity for delayed hearings on wait and sojourn times for all cases. While the two previously introduced policies can offer insights into avenues for reducing the size of the court backlog, these policies may serve cases too quickly and impact an individual's right to due process. What constitutes due process varies from case to case, particularly for cases involving persecuted and vulnerable persons. For example, extended time between hearings to secure legal counsel and prepare a case may be beneficial. In contrast, additional time may only prolong the immigration court process for another case. Court-caused delays represent added wait time to cases for reasons outside of the applicant's control and, in general, represent operational interruptions such as an absent judge or the prioritization of other cases. These factors often do not change the fairness of a case. Therefore, while measuring a case's due process is complex, deploying make-up capacity can reduce extra wait time without altering due process and is assessed in Section 5.5.

5.2.4. Baseline Model Modifications

Each policy implementation is a modified version of the baseline model presented in Chapter 4, Section 4.4. We adapt Model B, where judges are modeled using quartiles, which we henceforth refer to as the baseline model. The baseline model may be modified in two ways: i) input changes reflecting changes in the pre-defined data read into the model, such as service or arrival rates, or ii) function changes reflecting changes to the rules of the simulated environment, which are adjusted in the code functions. Table 5.2 depicts the modifications required for each policy implementation.

All experiments were run using SimPy 3.0.11 [302] and Python 3.7.3, with up to 128 GB of memory, under Linux Ubuntu 20.04.4 on Worcester Polytechnic Institute's high-performance computing research cluster. Each

Table 5.2: **Baseline Modification for Each Policy.**

	Input changes	Function changes
Policy 1: Varying the Number of New Judges	✓	
Policy 2: Dedicated dockets	✓	✓
Policy 3: Make-up Capacity		✓

policy was deployed starting in 2010 and was replicated 10 times. The runtime (wall clock time) for each replication ranged from 2 to 6 days. The methodology and results of each policy are presented in the subsequent sections.

Section 5.3

Policy 1 Implementation

We consider variations of Policy 1 as outlined in 7 scenarios in table 5.3.

Table 5.3: **Policy 1 Variation Descriptions.**

Policy 1 Variation	Description
Empirical Data	139 Judges
A	1 Judge Added
B	5 Judges Added
C	10 Judges Added
D	15 Judges Added
E	25 Judges Added
F	50 Judges Added
G	100 Judges Added

Each variation adds a total number of new judges to the system at varying levels (denoted as x in Table 5.5). Immigration judges have different behaviors in how frequently they serve cases and at what rates. Therefore to examine the impact of adding new judges to the system, it is necessary to assign the behaviors of the new judges in a way that is representative of the real-world setting.

Considering the baseline model groups similar judges together in quartiles based on average service rate (Chapter 4, Section 4.4), we leverage the observed frequency of each quartile and assign new judges proportionally. For each variation of Policy 1, the total number of new judges (n_i) assigned to each quartile is determined using the percentages shown in Table 5.4 and Equation 5.1a. The corresponding capacity of new judges is proportionally increased (p_i, Equation 5.1b). The total new capacity for each quartile on a given day is determined by adding the increased capacity to the existing capacity, as shown in Equation 5.1c. For example, suppose the original capacity assigned to a quartile on a given day is 100, and two new judges are added to the quartile. In that case, the proportional increase in capacity is 2%, and the new total capacity is 102. Additional daily capacity is rounded

Table 5.4: **The Number of Judges Added to Each Quartile across Policy 1 Variations.**

Model Variation	Number of Judges Added (n_i) Q1	Q2	Q3	Q4
Empirical Number of Judges (N_i)	21 Judges (15%)	8 Judges (6%)	47 Judges (34%)	63 Judges (45%)
A: 1 Judge Added	0	0	0	1
B: 5 Judges Added	1	0	2	2
C: 10 Judges Added	2	1	3	4
D: 15 Judges Added	2	1	5	7
E: 25 Judges Added	4	2	8	11
F: 50 Judges Added	8	3	17	22
G: 100 Judges Added	15	6	34	45

Table 5.5: **Policy 1 New Judge Capacity Assignment Variables.**

Symbol	Definition
x	The number of new judges to be added to the court
P_i	Percent of original judges in quartile i
p_i	Percent increase factor for judge quartile i
N_i	The original number of judges in quartile i
n_i	The number of new judges to be added to each quartile i as shown in Table 5.4
$C_{i,j}$	The original capacity of judge quartile i at day j
$C_{i,j}^{new}$	New capacity assigned to judge quartile i at day j

to the nearest integer.

Assigning new judges in a manner proportional to the observed quartiles maintains the variability of the baseline model by ensuring that new judges serve cases in quartiles with the highest demand while preserving the presence of less common judge behavior (quartiles with smaller percentages).

The deployment of adding new judges in the presented simulation model holds three assumptions: i) all physical resources and the associated costs of hiring new judges are trivial, ii) behaviors and the number of judges are constant for the duration of the simulated period and iii) a reduction in the interarrival times between hearings does not impact the due process of a case. While these assumptions are a limitation in the practicality of Policy 1, the described method, nevertheless, provides a robust simulation with inputs and outputs that are representative of the real-world setting to evaluate the impact of adding new judges to the immigration court system as described below.

$$n_i = \lfloor (P_i * x) + 0.5 \rfloor \quad (5.1a) \qquad p_i = \frac{n_i}{N_i} \quad (5.1b) \qquad C_{i,j}^{new} = \lfloor (C_{i,j} * p_i) + 0.5 \rfloor + C_{i,j} \quad (5.1c)$$

5.3.1. Policy 1 Results and Discussion

Figure 5.4 depicts the number of cases remaining in the system for all cases that are completed prior to the cutoff of the data time window (see Chapter 4, Section 4.3.1).

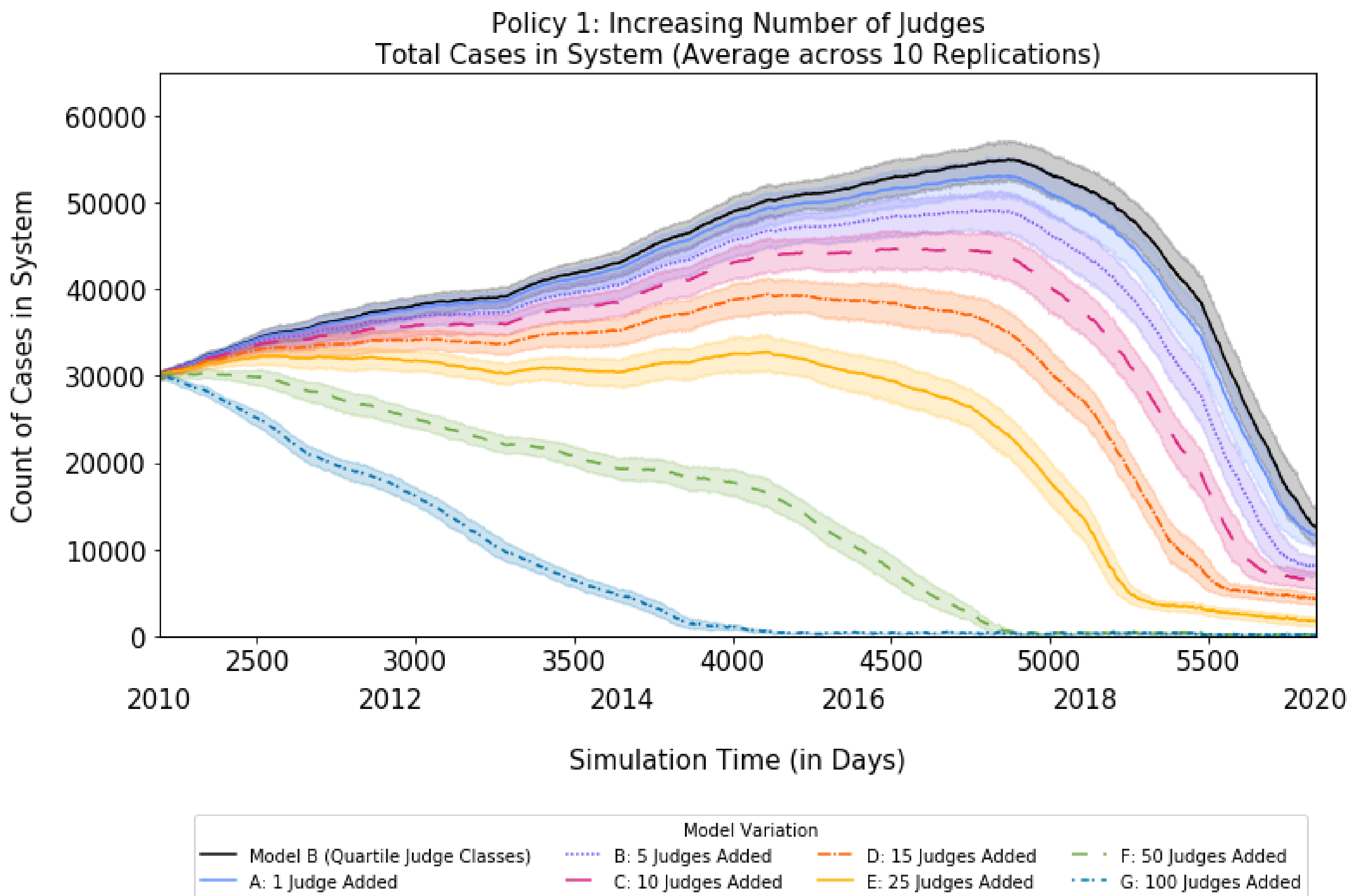

Figure 5.4: **Comparison of Cases Remaining in the System across Policy 1 Variations.**

It is clear that as the number of judges increases, the service rate also increases. The increased service rate means the number of cases remaining in the system has a faster convergence toward zero. While variations E, F, and G may propose a disproportional number of new judges for a single court (regarding resources and fairness, Section 5.2.1), these extreme instances help illustrate important aspects of the court backlog.

For example, under the hypothetical assumption that 100 new judges (variation G) are hired, the backlog of the NYC immigration court could be reduced by 50,000 cases (compared to the baseline model) in approximately five years, whereas hiring 50 (variation F) would take closer to seven years, and hiring 25 (variation E) over ten years. For each variation depicted in Figure 5.4, a peak “stress” value exists, depicting the maximum number of cases remaining in the system [4] during the simulated time. Figure 5.5 shows each variation's peak stress point value across the 10 replications. As the number of judges increases, the peak stress value and variance across replications decrease. The observed decrease reflects cases being served faster than cases are arriving. In the baseline model, the number of cases arriving exceeds those being served, causing the system to be unstable and the backlog unbounded. However, as shown in Figure 5.4, variations F and G both have decreasing trends and lower peak stress point values (Figure 5.5). Such observations indicate that the rate of cases exiting the system is greater than those arriving, and the system can stabilize to a steady state. While these instances demonstrate a rapid decrease in the size of the backlog, the utilization of judges also decreases, indicating a diminishing return on investment for hiring a larger number of new judges.

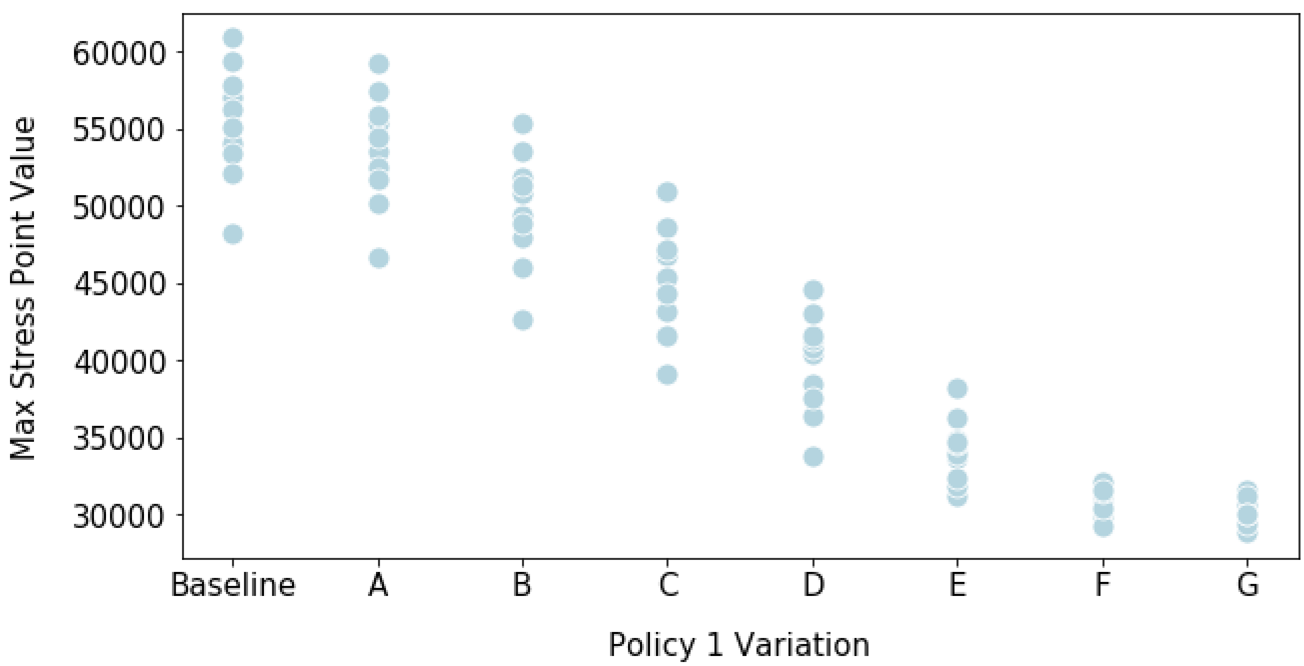

Figure 5.5: **Peak Stress Value Identified across Policy 1 Variations.**

While the previous Figures 5.4 and 5.5 highlight the impact judge hiring has on the size and rate of backlog reduction, they are limited to assessing the entire system rather than considering individual cases. Complementary Figures 5.6 and 5.7 examine the average reduction in mean wait and sojourn time for cases across each variation. Additional figures depicting the average reduction for other KPIs across each variation can be found in Appendix C.

The average reduction for the mean sojourn time (Figure 5.6 ranges from 18 days for variation A and close to 390 days for variation G. Figure 5.7 shows the average reduction for the mean wait time that each case experiences across hearing types. Overall, a proportional increase in the average reduction is observed as more judges are added

[4]The number of cases remaining in the system where the next hearing is known

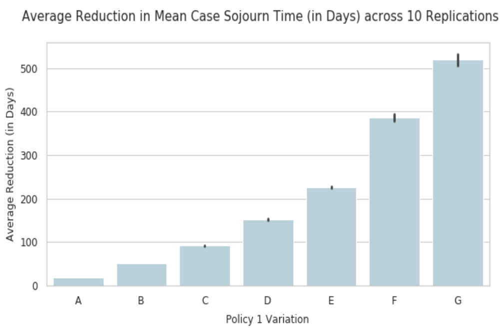

Figure 5.6: **Average Reduction in Mean Sojourn Time across Policy 1 Variations.**

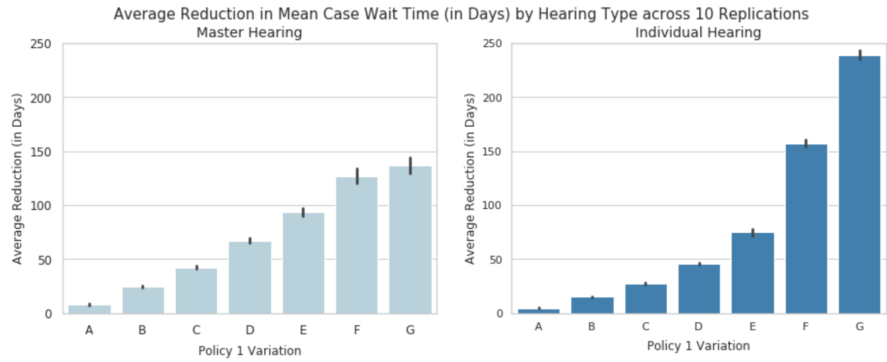

Figure 5.7: **Average Reduction in Mean Wait Time across Policy 1 Variations.**

across both hearing types, with the notable exception of variations F and G. For Master calendar hearings, these two variations produce a similar reduction, despite variation G adding double the number of judges. Variations F and G have much larger average reduction rates for Individual hearings than the other variations. Together these observations indicate that the benefit of hiring more judges has a larger impact on Individual hearing wait times compared to Master hearing wait times. This observation aligns with the real-world setting as Individual hearings tend to have less capacity each day, causing a bottleneck in the system. Therefore when more judges are available, the bottleneck is reduced, and cases are served faster. In contrast, Master hearings are shorter in nature and have much quicker service rates; therefore, additional judges appear to have a diminishing return on investment, as seen between variations F and G. While the largest total gain in reduction across all KPIs is for variation G, the rate of reduction has a decreasing trend as more judges are added[5].

The presented analytical assessment of Policy 1 depicts general trends and insights into the impact and "value-added" of hiring more judges to serve the growing backlog of cases in the NYC immigration court. Given the limited resources of the immigration court system, understanding the value and trade-offs related to hiring additional judges provides relevant planning information for decision makers. As expected, hiring 100 (variation G) new judges provides the largest reduction in wait and sojourn times. However, the additional reduction in time when the number of judges is only half that, that is, 50 (variation F), was minimal. The observed decreasing return on investment suggests a continual increase in judges may not have the intended impact. Furthermore, the practicality of hiring such large quantities of judges must be considered in terms of both operations and ethics, although these aspects were not captured in the presented model. The temporary hiring of additional judges may be an avenue for future exploration as immigration judges do not hold life appointments [295]; however, the physical resources (and associated costs) to support a rapid addition of judges may limit the practicality. It should be noted that the proposed Policy 1 only measures the rate of service when hiring more judges and does not examine the potential

[5] This effect may partially be due to the variability of judge assignments to the different quartiles and requires further investigation.

loss in the due process associated with serving cases at a faster rate (a shorter interarrival time between hearings). When considering policy implications, future work should explore the various trade-offs between efficiency, fairness and equity, and costs and seek to balance these in a data-informed manner. The presented representation of hiring more immigration judges forms a rich foundation for future variations of Policy 1.

Section 5.4

Policy 2: Dedicated Dockets

We implement the policy of dedicated dockets by creating new queues and reserving varying amounts of capacity for asylum cases. It is assumed that for each judge and hearing type, a dedicated docket is created and that judges regularly serve from both the dedicated and non-dedicated dockets. The service rate for each respective docket must be determined. Policy 2 seeks to shift capacity between asylum and non-asylum cases rather than adding new capacity as in Policy 1. In light of this, each docket type is assigned a proportion of the current capacity. Dedicated dockets are strictly used for asylum cases, and non-dedicated dockets are strictly used for non-asylum cases. Therefore, if all of the cases waiting in a particular docket are served, and extra capacity exists, it goes unused, even when cases remain in the other docket. We test the sensitivity of different capacity assignment schemes (variations) as outlined in Table 5.6.

Table 5.6: **Policy 2 Variation Descriptions.**

Policy 2 Variation	Description
Empirical Average Splits	Master Hearings: Asylum: 50%, Non-Asylum: 50% Individual Hearings: Asylum: 77%, Non-Asylum: 23%
A	Average Splits by Year
B	Average Splits by Judge
C	Average Splits by Judge, Year
D	25% Asylum, 75% Non-Asylum
E	75% Asylum, 25% Non-Asylum
F	50% Asylum, 50% Non-Asylum
G	Average Splits by Presidential Administration
H	Increase Asylum over Empirical Splits by 5%; Decrease Non-Asylum by 5%
I	Decrease Asylum under Empirical Splits by 5%; Increase Non-Asylum by 5%
J	Increase Asylum over Empirical Splits by 5%; Decrease Non-Asylum by 2.5%
K	Decrease Asylum under Empirical Splits by 5%; Increase Non-Asylum by 2.5%

The observed splits for asylum versus non-Asylum cases, averaged in the empirical data across 2010-2019 for all judges in NYC, are 50%, 50% for the Master calendar hearing, and 77% and 23% for the Individual hearing. Variations A, B, C, and G assign the capacity for dedicated dockets based on four generalized behaviors: year, judge, judge and year, and presidential administration. The remaining variations apply the same splits to both Master and Individual hearings. Variations D, E, and F are more aggressive splits and prescribe the total percentage of capacity assigned to each docket. Variations H, I, J, and K shift the empirical daily capacity between dedicated and non-dedicated dockets at rates of 2.5% and 5%. We use dedicated docket and asylum cases and non-dedicated docket and non-asylum cases interchangeably where unambiguous.

5.4.1. Policy 2 Results and Discussion

The total number of cases remaining in the system from 2010-2019 across the 11 variations are shown in Figure 5.8. Variations D, E, and F show a large growth in the size of the immigration court backlog. Dedicated dockets represent a shift in how capacity is used rather than an increase or decrease in available capacity. Therefore, the system's expected behavior is to follow more closely to the baseline model as observed across the other variations. The explosive behavior observed in the number of cases under variations D, E, and F results from underutilizing capacity. In other words, under these variations, too much capacity is allocated to either asylum or non-asylum cases. Because the capacity cannot be shared between dockets, some excess capacity goes unused, leading to more cases in the system.

Figure 5.9 shows the average time a case waits at each hearing type for asylum and non-asylum cases, and Figure 5.10 shows the average sojourn time. The largest changes for both KPIs occur in variations D, E, and F because, as previously mentioned, these variations have the most extreme deviation in capacity assignments.

Variation G produces results nearly identical to the baseline model across all KPIs, indicating the average splits observed across administrations capture the overall behavior of the baseline model well. While additional exploration is required, this indicates that the presidential administration greatly influences how many asylum and non-asylum cases are processed over time. Often the policies and memoranda put into place by different administrations impact how cases enter and interact with the system and further support this observation (see Sections 4.2.5 and 5.2.2). Variations H, I, J, and K propose conservative shifts in capacity between asylum and non-asylum cases. Variations H (5%) and J (2.5%) both seek to shift *more* capacity towards serving the dedicated docket, whereas I (5%) and K (2.5%) seek to shift capacity towards serving the non-dedicated docket. Comparing average sojourn time, the effects of increasing capacity for asylum cases (variations H and J) are greater than non-asylum cases (variations I and K). Notably, variation J results in (almost) equal average sojourn times for asylum and non-asylum cases. However, this equality in sojourn times comes at the cost of, on average, longer sojourn and wait times for non-asylum cases. If the goal is to have more equal average wait times between asylum and non-asylum cases at each hearing, then the status quo (baseline model) is preferred. However, the baseline model has longer sojourn times for asylum cases. It should be noted that longer sojourn times in the empirical data are sometimes attributed to applicant delays, where the case required more time to prepare for their hearing

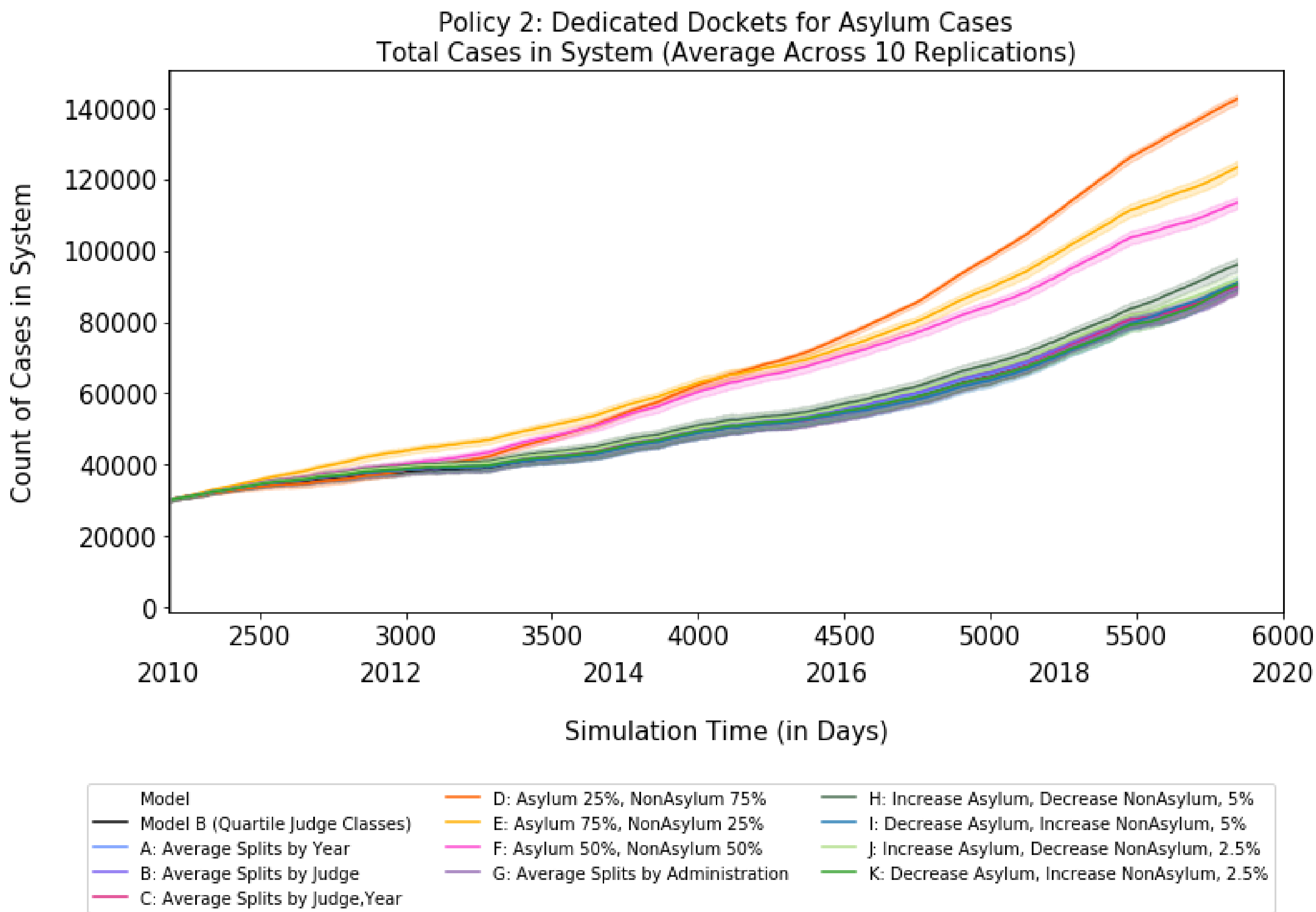

Figure 5.8: **Comparison of Cases Remaining in the System across Policy 2 Variations.**

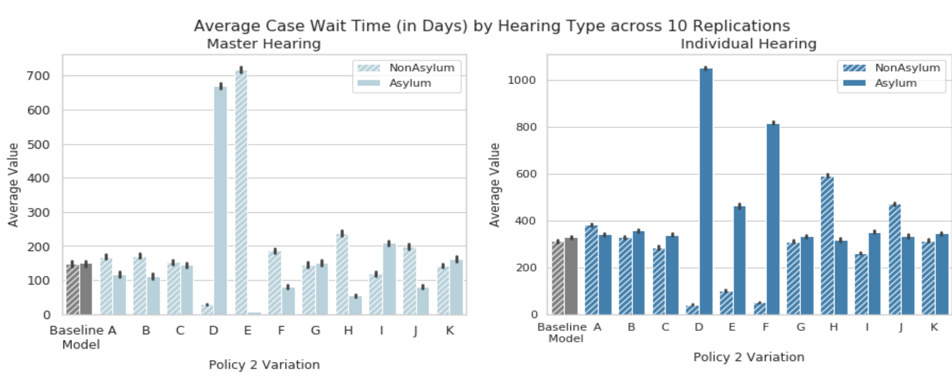

Figure 5.9: **Average Case Wait Time across Policy 2 Variations.**

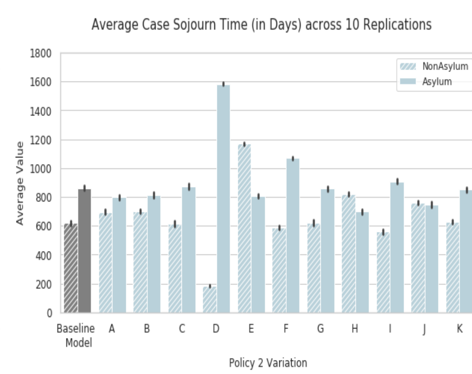

Figure 5.10: **Average Sojourn Time across Policy 2 Variations.**

and may be indicative of due process.

Policy 2 examines the impact of using dedicated dockets and shifting resources for asylum cases at varying levels providing insights regarding case equity and trade-offs in average sojourn and wait times. More aggressive changes in serving either dedicated or non-dedicated dockets resulted in increasing the size of the backlog as capacity is not shared between the dockets and therefore went unused. Non-asylum cases experience a larger increase in wait times at each hearing type when additional capacity is reserved for asylum cases, suggesting non-asylum cases are more sensitive to shifts in capacity. However, shifting capacity in relatively small amounts towards serving more asylum cases (variation J) was found to produce more balanced sojourn times across asylum and non-asylum cases. Under Variation J the sojourn time for both types of cases is, on average, close to 750 days – a decrease of about 114 days for asylum cases and an increase of about 140 days for non-asylum cases. Considering due process is difficult to measure, the fairness of serving cases at expedited speeds remains an area of future work.

Section 5.5

Policy 3: Make-up Days

Policy 3 seeks to prioritize the serving of cases that experience a court-caused delay by reserving "make-up capacity" for these hearings in semi-regular intervals. It is assumed that each judge and hearing type reserve make-up capacity in each interval. We consider two variations for deploying make-up capacity on reserved days as described in Table 5.7.

Table 5.7: **Policy 3 Variation Descriptions.**

Policy 3 Variation	Description
A	Capacity Reserved Every 7 Days
B	Capacity Reserved Every 14 Days

For each variation, we examine the effects make-up capacity has on the court backlog size and the average wait and sojourn times for cases with delayed and non-delayed hearings. Using the same service rates as the baseline model, capacity is reserved to allow for the court-caused delayed hearings to be "made up" every 7 or 14 days.

The objective of Policy 3 is to reduce the wait time added due to court-caused delays by reserving make-up capacity in semi-regular intervals, called "reserved days". On reserved days, cases with the shortest anticipated delay are served first, and for occurrences where multiple entities have the same anticipated delay, cases are served first-come, first-serve (FCFS). While a hearing may be delayed for a plethora of reasons, it is assumed that all court-caused delays have equal priority within their class. In some instances, the demand for make-up capacity in a given interval may exceed (or be less than) the amount available. Unlike dedicated dockets (Policy 2), where capacity goes unused given that two separate queues exist for asylum and non-asylum cases, delayed and non-delayed hearings are served from the same queue. Therefore, capacity can be shared between delayed and non-delayed hearings. To ensure the optimal utilization of capacity, a set of service rules are established.

First, if a case is seeking service on a reserved day but no more capacity remains, the case waits until the next reserved day. Should a delayed case have an opportunity to be served before the next available reserved day, the case is seen by a judge. Finally, if make-up capacity remains after all delayed cases are served, non-delayed cases are served. By creating such rules, we present a policy that reduces the underutilization of capacity, ensuring a favorable presentation of the proposed policy.

Implementing this policy in the simulated environment requires daily checking of the status of entities with a delay until service is received. This process is computationally and memory-intensive. For these reasons, Policy 3 is deployed for only five years (2010-2015) across 10 replications. Considering make-up capacity is intended to improve the short-term reduction in hearing wait times, five years is deemed adequate to capture the intended impact.

5.5.1. Policy 3 Results and Discussion

Figure 5.11 depicts the total number of cases in the system from 2010 to 2015 for each variation of Policy 3. Variations A and B follow the same general trend as the baseline model, with, on average, a daily increase of 4.0% and 2.5% in total cases in the system, respectively. The cause for the backlog increase is a result of fewer cases exiting and is examined in what follows.

Figure 5.12 shows the average wait time for delayed and non-delayed hearings, and Figure 5.13 compares the average sojourn time for cases with and without a delay. Figure 5.14 more closely depicts the average *reduction* in case wait times for delayed and non-delayed hearings. The x-axis depicts the average reduction in days, where negative numbers represent an increase in the mean wait time. Variations A and B result in reductions for delayed cases, with only a relatively small increase (that is, negative reduction) for non-delayed cases at both Master and Individual hearings. Variation A reserves capacity more frequently (every 7 days), and consequently, a greater benefit for delayed cases are incurred when compared to variation B (every 14 days). The wait time associated with delayed cases is reduced by approximately 122 (79) days and 162 (23) days at a Master hearing in variation A (B) and Individual hearing, respectively.

In contrast to the reduction in wait times for delayed cases, under Policy 3, the implementation of variation B resulted in an increase in total sojourn time for cases with at least one delay (Figure 5.13). Investigating further the total number of completed cases (Figure 5.15), we see a decrease for cases with and without at least one delay. Thus, the decrease in completed cases accounts for the increase in the total cases *remaining* in the system. However, the reason for the decrease in completed cases requires additional exploration. Considering sojourn time represents the total time spent in the system, examining the number of cases served for Master and Individual hearings may provide additional insights.

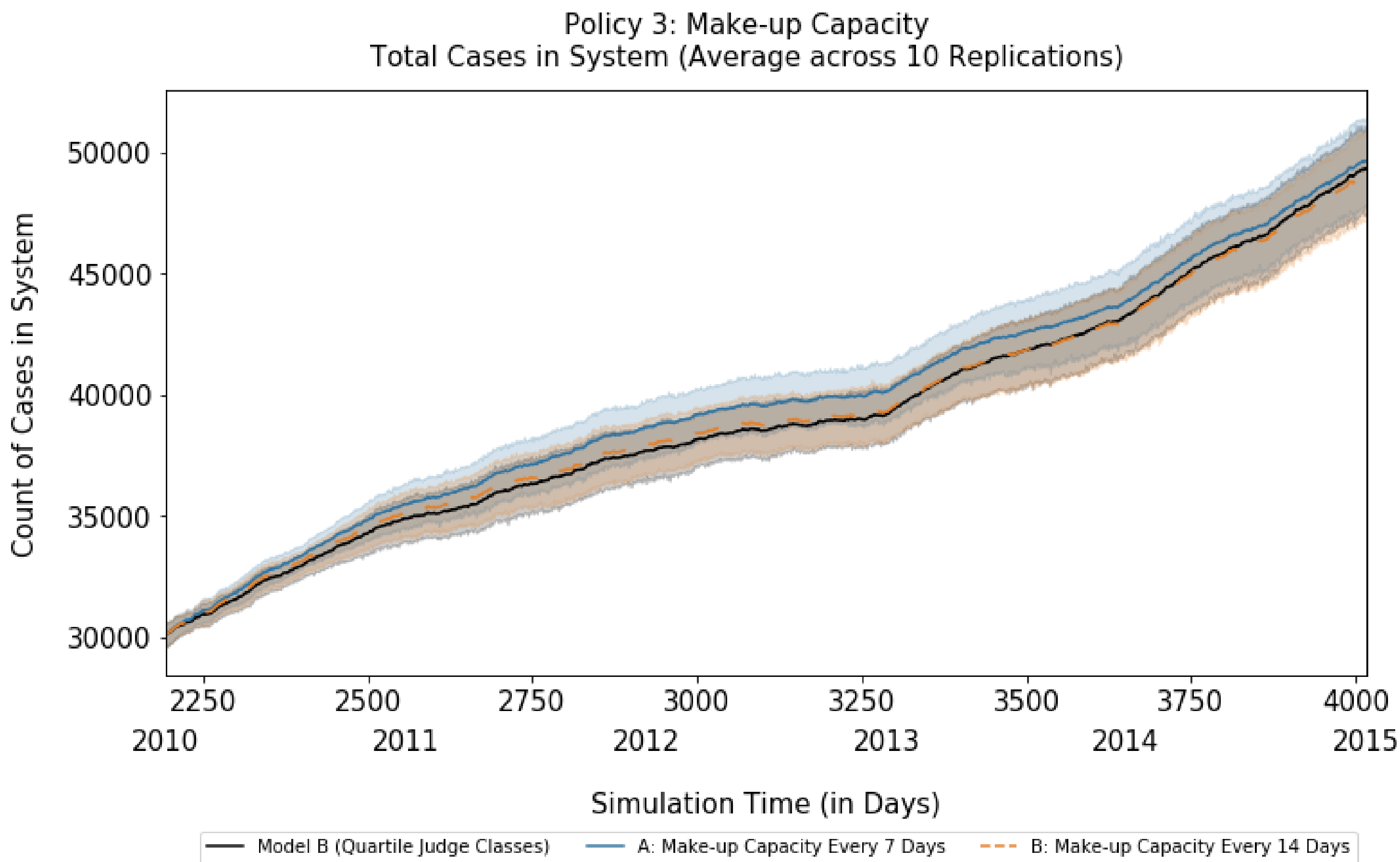

Figure 5.11: **Comparison of Cases Remaining in the System across Policy 3 Variations.**

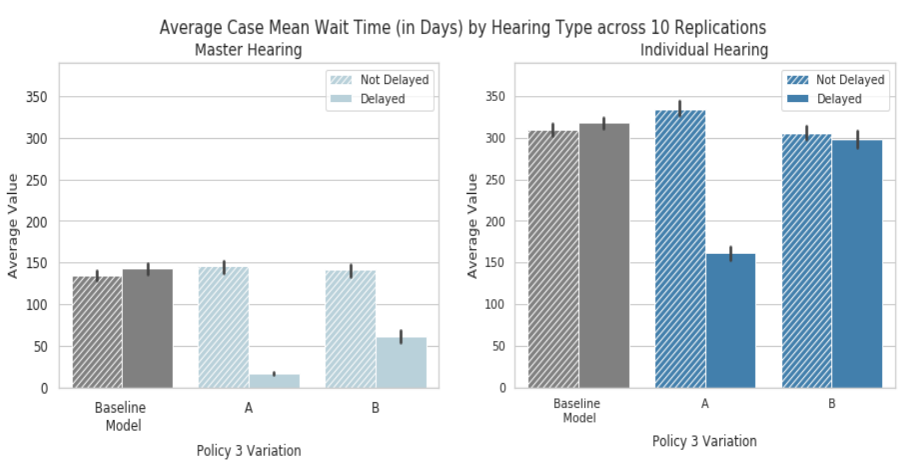

Figure 5.12: **Average Hearing Wait Time (in Days) across Policy 3 Variations.**

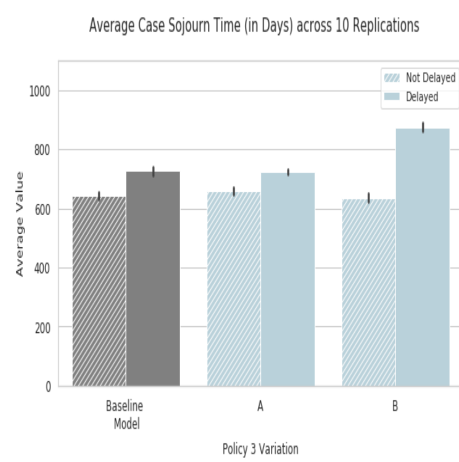

Figure 5.13: **Average Case Sojourn Time (in Days) across Policy 3 Variations.**

Figure 5.16 shows the average decrease in the number of delayed and non-delayed hearings served at each hearing type. In variation A, on average, 1,773 additional delayed cases were served at the Master hearing and over 10,000 at the Individual hearing. The large shift in serving cases with a delay in the Individual hearing is likely

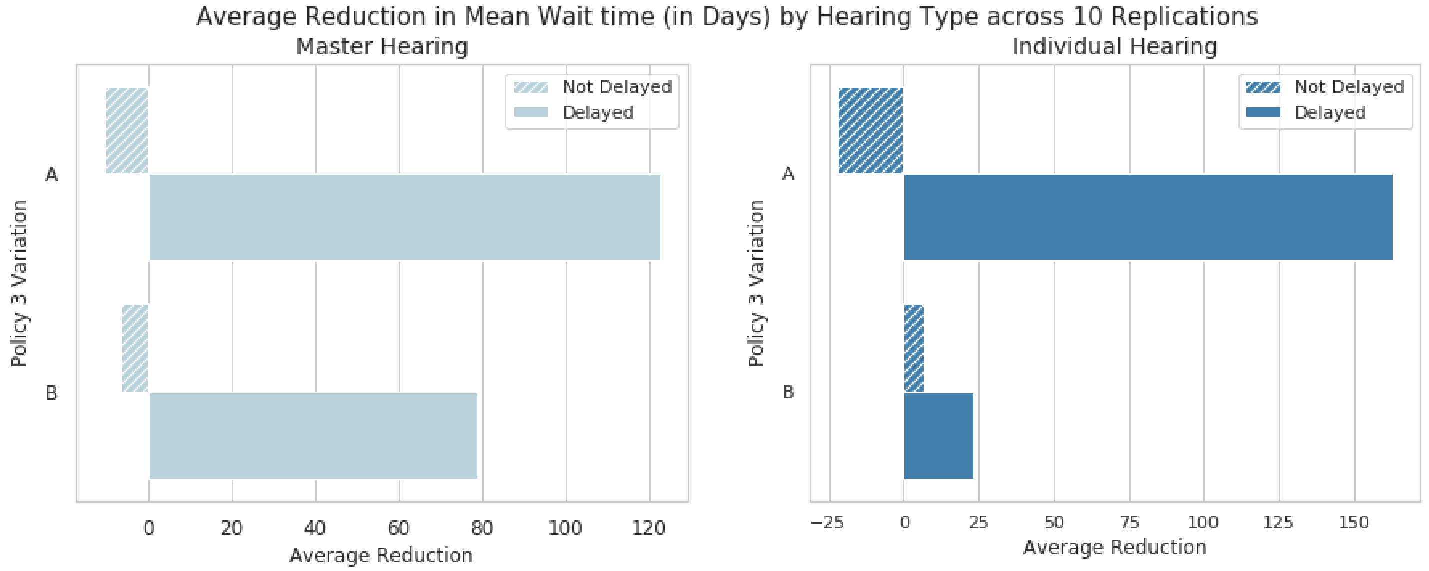

Figure 5.14: **Average Hearing Wait Time (in Days) across Policy 3 Variations.**

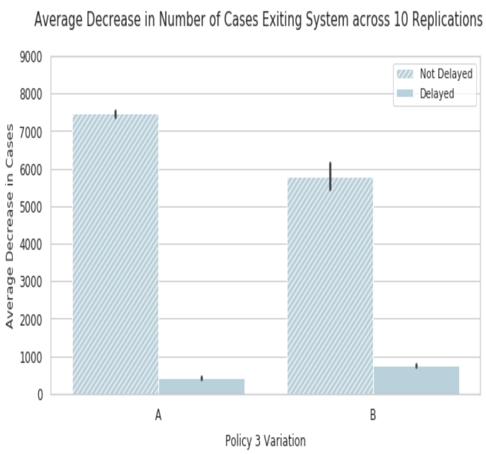

Figure 5.15: **Average Decrease in Cases Exiting the System across Policy 3 Variations.**

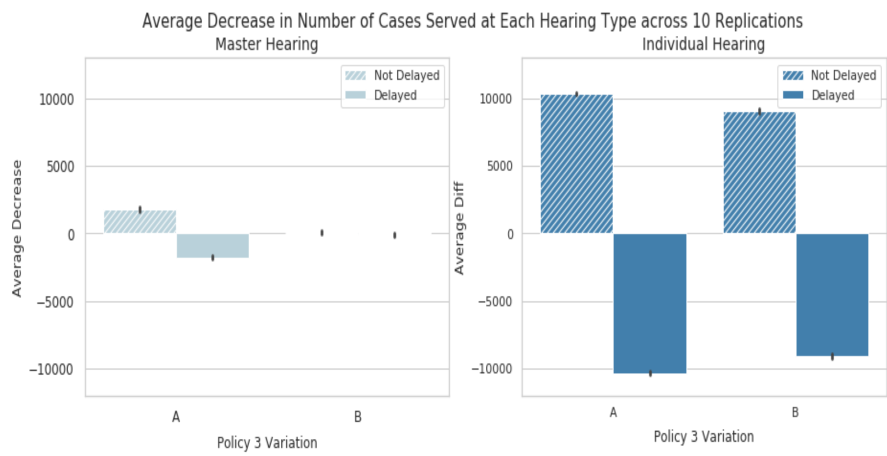

Figure 5.16: **Average Decrease in Cases Exiting Each Hearing Type across Policy 3 Variations.**

a significant contributing factor to the decrease in total cases exiting the system. In the empirical data, delays at an Individual hearing occur approximately 30% of the time (Appendix C, Figure 27). In addition, cases with delays tend to have longer sequences and experience more than one delay (Appendix C, Figure 30). Therefore, while make-up capacity reserved for delayed cases helps to reduce their average wait time at the Individual hearing, the propensity of delayed cases to experience multiple delays and have longer sequences results in a greater number of cases lingering in the system, with fewer exiting. The treatment of delayed vs. non-delayed hearings is the same in the simulation model; delayed hearings are considered a step in the sequence rather than strictly time added (Appendix C, Figure 29). Therefore, delayed hearings are assigned an anticipated delay (place in the queue, see Section 4.3.2) and must wait until capacity is available and "service" received before moving on to the next hearing in their sequence. Accordingly, modeling delays in this way may further contribute to the bottleneck observed at the Individual hearing.

In an effort to reduce system KPIs, Policy 3 deploys make-up capacity to reallocate capacity for delayed hearings and prioritize them on reserved days. Reductions in the average wait time for delayed cases were observed and represent the short-term gains of make-up capacity. However, long-term Policy 3 further aggregated the existing bottleneck for Individual hearings and resulted in fewer cases exiting the system, leading to an increase in the number of cases remaining in the system. Nevertheless, the overall decrease in wait times for delayed cases at little cost to non-delayed cases suggests the deployment of make-up capacity may be a viable option to reduce extra wait time without altering the due process of cases. In particular, the results of this policy suggest that the deployment of make-up capacity may best be suited for Master calendar hearings.

Policy 3 was tested within the parameters of our baseline model assumptions. While make-up capacity was not fully effective in reducing case wait and sojourn times, valuable insights were gleaned. Future work can build upon the presented foundation to test and evaluate future variations of Policy 3, such as i) larger make-up capacity intervals (e.g., 21 and 30 days), ii) deployment of make-up capacity only for Master calendar hearings, and iii) using prioritization between delay types.

Section 5.6

Conclusion

This chapter presents the assessment of three policies, each with several variations. The policies were informed by our analysis and supported by domain experts.

Policy 1 compared the impact of hiring a varying number of judges. While the variation with the largest number of judges reduced the backlog at the greatest rate, the results indicated the "value-added" of judges might have a diminishing return on investment. Furthermore, the practicality of hiring such large quantities of judges must be considered in terms of operational realities, including costs, as well as ethics, such as due process, because these aspects were not captured in the presented model.

Policy 2 evaluated the impact that dedicated dockets for asylum cases have on overall system behaviors. The preferred amount of judge capacity allocated toward serving dedicated dockets differs, depending on the objective. Allocating large amounts of capacity towards either dedicated or non-dedicated dockets resulted in large increases in the size of the backlog – unsurprisingly, as capacity is unable to be shared between dedicated dockets, this results in unused capacity. This suggests that smaller shifts in capacity may produce more equitable results. For example, shifting 2.5% additional capacity (variation J) toward asylum cases resulted in almost equal sojourn times for all cases. However, this resulted, on average, in longer wait times for non-asylum cases at both the Master and Individual hearings. One limitation of the presented deployment of dedicated dockets is that capacity is assumed unable to be shared between asylum and non-asylum cases, and therefore, in a number of variations, results in underutilized capacity. In addition, judges are assumed to serve both asylum and non-asylum cases. Future work could explore sharing capacity between case types and assigning judges only to serve certain case types (e.g., asylum and non-asylum).

Policy 3 evaluates the novel deployment of make-up capacity for the purpose of reducing the additional wait time resulting from court-caused delays. The increased opportunity to serve delayed cases at the Master hearing resulted in an average wait time reduction of approximately 87%, while delayed cases at the Individual hearing resulted in a reduction of approximately 51%. While the reduction in wait time at both delays is welcomed, the prioritization of delayed cases at the Individual hearing contributed to an increase in the total number of cases in the system, suggesting make-up capacity may be best suited for the Master calendar hearing. While make-up capacity was not fully effective in reducing case wait and sojourn times, the process revealed valuable insights. Future work can build upon the presented foundation to test and evaluate the deployment of make-up capacity for only Master hearings across longer intervals or using prioritization between delay types.

The baseline model is for the NYC immigration court system, and care should be used before generalizing results to other court contexts. Rather, the specifics of other court systems can be used as input to this model to identify individual system performance across KPIs. Therefore, the baseline model can be applied to other immigration court locations to capture variation in behaviors, further enabling the development of a multi-location model. A multi-location model can provide an expansive representation of the immigration court system, capturing the interaction between different courts and supporting the evaluation of policy effectiveness for both individual courts and the system as a whole.

In addition, the baseline models infrastructure can be adapted to model other immigration-related processes, such as the affirmative asylum process handled by the United States Citizenship and Immigration Services (USCIS) [259]. While the USCIS affirmative asylum process is handled in a non-courtroom setting, there exist interactions between the USCIS and the EOIR. For example, if an affirmative asylum application is denied, the individual(s) are subsequently handled in the EOIR, where some individuals seek asylum through the defensive process [315]. Opportunities exist to explore the redistribution of immigration-related responsibilities between these two agencies and are avenues for future work.

Changes to processes in the immigration court may impact millions of individuals. Undoubtedly great care is required with any system adjustments. The presented model demonstrates an analytical approach to evaluate policies for possible recommendation and future deployment in real-world contexts. The results presented illustrate the promising potential of using discrete event simulation as a decision-making mechanism to support the evaluation of immigration policy. The observed insights both provide foundational work and require further scrutiny in advance of making final recommendations. The depth of insights gleaned points to the great future promise of generating data-informed recommendations to improve the immigration court system. Future work should explore methods to measure the due process of cases under each scheme to provide important insights into their respective fairness. This analysis lays the foundations for an adaptable and interactive decision-making tool to support decision makers seeking efficient and more equitable solutions to the growing immigration court backlog and to improve operations for those interacting with the immigration court system. Collaboration with stakeholders and those working closely with the immigration court can improve the reliability and practicality of the presented model and is strongly cautioned prior to the deployment of any of the presented results.

Conclusion

This dissertation contributes to the *Data Science for Social Good* movement through the application of analytics to improve the operations of governmental agencies, non-governmental organizations (NGOs) and nonprofit organizations serving vulnerable populations in two important domains: anti-human trafficking and immigration. Notably, I demonstrate the power and value of insights gleaned from data-driven methods for decision makers in these two contexts.

In Chapter 1 we present an in-depth review of the current research landscape of Operations Research and Applied Analytics in the anti-human trafficking domain. The review highlights the current work being done and directions for future research. Chapter 2 contributes to the growing body of literature examined in Chapter 1 and analyzes data using Data Envelopment Analysis to evaluate the performance of border stations of NGO Love Justice International (LJI) engaged in anti-trafficking efforts. The results of the work provide an improved understanding of the current performance of different border stations and provided operational improvement recommendations for the organization's decision makers.

Chapters 3 - 5 discuss an analytical approach to model and assess the United States Immigration court system. Chapter 3 created a framework of the immigration court system and structure through the use of Queueing theory. Chapter 4 builds upon the previous queueing model framework and deploys a large-scale discrete event simulation model simulating the New York City immigration court system from 2004-2019. A baseline model is validated over 15 years of empirical data. The applicability of the model is demonstrated through sensitivity analysis and proffers insight into how changes in arrival and service rates affect key performance indicators, laying the groundwork for the work presented in the final chapter of this work.

Chapter 5 evaluates three policies, informed by our analysis and supported by domain experts. The influence of each policy is evaluated using three metrics that evaluate the reduction of sojourn times, wait times, and queue lengths. The testing of such policies within my model demonstrates the ability to generate data-informed insights for decision makers, something in critically short supply in this important aspect of our society. The three examined policies are: i) vary the number of new judges, ii) dedicated dockets for asylum cases and iii) make-up capacity for delayed cases. The first policy varies the total number of judges to illustrate the impact of the quantity of judges on system throughput. The second policy seeks to serve asylum cases more expeditiously using dedicated dockets. The third policy seeks to reduce wait time added due to court-cause delays through the introduction of "make-up" capacity. The implementation and evaluation of these three policies provide case-level insights for decision makers and illustrate the power of data-driven Discrete Event Simulation modeling. The framework of the developed baseline model provides abundant opportunities for the modeling and evaluation of additional policies as discussed in Sections 5.3 - 5.5.

Data Science and Applied Analytics are powerful tools that hold immense promise for governmental agencies, non-governmental organizations (NGOs) and nonprofit organizations. By leveraging data-driven methods, these organizations can operationalize their data to provide valuable insights for more informed decision-making. Through informed decision-making, governmental agencies, NGOs, and nonprofit organizations can improve their operations and utilization of resources and more effectively serve the needs of the most vulnerable in our society. As demonstrated in this work, in the anti-human trafficking and immigration domains, data-driven methods aid in the understanding of operations, identification of patterns, and improve resource allocation. Together these insights can be used to inform and evaluate policies. Beyond these specific domains, the applications of Data Science and Applied Analytics are vast and promising.

Bibliography

[1] Corbett-Davies S, Pierson E, Feller A, Goel S, Huq A. Algorithmic decision making and the cost of fairness. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery & Data Mining; 2017. p. 797–806.

[2] Kasy M, Abebe R. Fairness, equality, and power in algorithmic decision-making. In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency; 2021. p. 576–586.

[3] Ahani N, Andersson T, Martinello A, Teytelboym A, Trapp AC. Placement optimization in refugee resettlement. Operations Research. 2021;69(5):1468–1486. doi:https://doi.org/10.1287/opre.2020.2093.

[4] Ahani N, Farajzadeh F, zaltın O, Trapp AC. Risk-Averse Placement Optimization in Refugee Resettlement. Available at SSRN 4344080. 2023;.

[5] Perspectives on how to conduct responsible anti-human trafficking research in operations and analytics. European Journal of Operational Research. 2023;309(1):319–329. doi:https://doi.org/10.1016/j.ejor.2022.12.028.

[6] Dimas GL, Konrad RA, Lee Maass K, Trapp AC. Operations research and analytics to combat human trafficking: A systematic review of academic literature. PLoS one. 2022;17(8):e0273708. doi:https://doi.org/10.1371/journal.pone.0273708.

[7] Stylianou A, Schreier J, Souvenir R, Pless R. Traffickcam: Crowdsourced and computer vision based approaches to fighting sex trafficking. In: 2017 IEEE Applied Imagery Pattern Recognition Workshop (AIPR). IEEE; 2017. p. 1–8.

[8] Saxena D, Badillo-Urquiola K, Wisniewski PJ, Guha S. A human-centered review of algorithms used within the US child welfare system. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems; 2020. p. 1–15.

[9] Kaya YB, Maass KL, Dimas GL, Konrad R, Trapp AC, Dank M. Improving access to housing and supportive services for runaway and homeless youth: Reducing vulnerability to human trafficking in New York City. IISE Transactions. 2022; p. 1–15.

[10] Office to Monitor and Combat Trafficking in Persons. 2019 Trafficking in Persons Report. United States Department of State; 2019. Available from: https://www.state.gov/reports/2019-trafficking-in-persons-report/.

[11] United Nations. Protocol to prevent, suppress and punish trafficking in persons; 2000. Available from: www.ohchr.org/en/professionalinterest/pages/protocoltraffickinginpersons.aspx.

[12] International Labour Organization. Global estimates of modern slavery: Forced labour and forced marriage; 2017. Available from: https://www.ilo.org/global/publications/books/WCMS_575479/lang--en/index.htm.

[13] International Labour Office. Profits and poverty: The economics of forced labour; 2014. Available from: https://www.ilo.org/global/topics/forced-labour/publications/profits-of-forced-labour-2014/lang--en/index.htm.

[14] Gallo M, Konrad RA, Thinyane H. An Epidemiological Perspective on Labor Trafficking. Journal of Human Trafficking. 2022;8(2):113–134. doi:https://doi.org/10.1080/23322705.2020.1815474.

[15] Anthony B, Penrose JK, Jakiel S, Couture T, Crowe S, Fowler M, et al. The typology of modern slavery: Defining sex and labor trafficking in the United States. Washington, DC, USA: Polaris. 2017;.

[16] Amadasun S. Social work interventions for human trafficking victims' in Nigeria. International Social Work. 2020;1:1–13. doi:https://doi.org/10.1177/0020872819901163.

[17] Okech D, Choi YJ, Elkins J, Burns AC. Seventeen years of human trafficking research in social work: A review of the literature. Journal of Evidence-Informed Social Work. 2018;15(2):103–122. doi:https://doi.org/10.1080/23761407.2017.1415177.

[18] Potocky M. The travesty of human trafficking: A decade of failed US policy. Social Work. 2010;55(4):373–375. doi:https://doi.org/10.1093/sw/55.4.373.

[19] Fraley HE, Aronowitz T, Stoklosa HM. Systematic review of human trafficking educational interventions for health care providers. Western Journal of Nursing Research. 2020;42(2):131–142. doi:https://doi.org/10.1177/0193945919837366.

[20] Hemmings S, Jakobowitz S, Abas M, Bick D, Howard LM, Stanley N, et al. Responding to the health needs of survivors of human trafficking: A systematic review. BMC Health Services Research. 2016;16(1):1–9. doi:https://doi.org/10.1186/s12913-016-1538-8.

[21] McCarthy LA. Trafficking Justice. Cornell University Press; 2016. Available from: https://www.cornellpress.cornell.edu/book/9781501701368/trafficking-justice/#bookTabs=1.

[22] Farrell A, Owens C, McDevitt J. New laws but few cases: Understanding the challenges to the investigation and prosecution of human trafficking cases. Crime, Law and Social Change. 2014;61(2):139–168. doi:https://doi.org/10.1007/s10611-013-9442-1.

[23] Farrell A, Fahy S. The problem of human trafficking in the US: Public frames and policy responses. Journal of Criminal Justice. 2009;37(6):617–626. doi:https://doi.org/10.1016/j.jcrimjus.2009.09.010.

[24] De Hert P, Muraszkiewicz J. Gary Becker and the economics of trafficking in human beings. New Journal of European Criminal Law. 2014;5(2):116–120. doi:https://doi.org/10.1177/203228441400500201.

[25] Mahmoud TO, Trebesch C. The economics of human trafficking and labour migration: Micro-evidence from Eastern Europe. Journal of Comparative Economics. 2010;38(2):173–188. doi:https://doi.org/10.1016/j.jce.2010.02.001.

[26] McAlpine A, Kiss L, Zimmerman C, Chalabi Z. Agent-based modeling for migration and modern slavery research: A systematic review. Journal of Computational Social Science. 2021;4(1):243–332. doi:https://doi.org/10.1007/s42001-020-00076-7.

[27] Wen J, Klarin A, Goh E, Aston J. A systematic review of the sex trafficking-related literature: Lessons for tourism and hospitality research. Journal of Hospitality and Tourism Management. 2020;45:370–376. doi:https://doi.org/10.1016/j.jhtm.2020.06.001.

[28] Mahalingam R. Human trafficking from a multidisciplinary perspectives: A literature review. Asian Journal of Social Science Research. 2019;2(2).

[29] Szablewska N, Kubacki K. Anti-human trafficking campaigns: A systematic literature review. Social Marketing Quarterly. 2018;24(2):104–122. doi:https://doi.org/10.1177/1524500418771611.

[30] Schauer EJ, Wheaton EM. Sex trafficking into the United States: A literature review. Criminal Justice Review. 2006;31(2):146–169. doi:https://doi.org/10.1177/0734016806290136.

[31] Konrad RA, Trapp AC, Palmbach TM, Blom JS. Overcoming human trafficking via operations research and analytics: Opportunities for methods, models, and applications. European Journal of Operational Research. 2017;259(2):733–745. doi:https://doi.org/10.1016/j.ejor.2016.10.049.

[32] Kammer-Kerwick M, Busch-Armendariz N, Talley M. Disrupting illicit supply networks: New applications of operations research and data analytics to end modern slavery. Bureau of Business Research; 2018. Available from: https://repositories.lib.utexas.edu/bitstream/handle/2152/68149/kammer-kerwick-2018-nsf-disrupting-illicit-supply-networks.pdf?sequence=2.

[33] Caulkins JP, Kammer-Kerwick M, Konrad R, Maass KL, Martin L, Sharkey T. A call to the engineering community to address human trafficking. National Academy of Engineering – The Bridge. 2019;49(3):68–73. doi:http://dx.doi.org/10.26153/tsw/3207.

[34] Office to Monitor and Combat Trafficking in Persons . The 3Ps: Prevention, Protection, Prosecution; 2016. Online. Available from: https://2009-2017.state.gov/documents/organization/259311.pdf.

[35] Moher D TJAD Liberati A. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA statement. PLoS Med. 2009;6(7):1–6.

[36] Raets S, Janssens J. Trafficking and technology: Exploring the role of digital communication technologies in the Belgian human trafficking business. European Journal on Criminal Policy and Research. 2021;27(2):215–238. doi:https://doi.org/10.1007/s10610-019-09429-z.

[37] Farrell A, De Vries I. Measuring the nature and prevalence of human trafficking. The Palgrave international handbook of human trafficking. 2020; p. 147–162. doi:https://doi.org/10.1007/978-3-319-63058-8_6.

[38] Weitzer R. New directions in research on human trafficking. The ANNALS of the American Academy of Political and Social Science. 2014;653(1):6–24. doi:https://doi.org/10.1177/0002716214521562.

[39] Acharjee S, Sarma DJ, Hanneman RA, Mordeson JN, Malik DS. Fuzzy soft attribute correlation coefficient and application to data of human trafficking. Proyecciones (Antofagasta). 2018;37(4):637–681. doi:https://doi.org/10.4067%2Fs0716-09172018000400637.

[40] Adam I, Al Qunaibit A, Shabeb N, Al Fehaid F, et al. Web application based image geolocation analysis to detect human trafficking. Journal of Information Security and Cybercrimes Research. 2021;4(2):185–193.

[41] Akram M, Zafar F. A new approach to compute measures of connectivity in rough fuzzy network models. Journal of Intelligent & Fuzzy Systems. 2019;36(1):449–465. doi:https://doi.org/10.3233%2Fjifs-181751.

[42] Ali S, Mathew S, Mordeson JN. Hamiltonian fuzzy graphs with application to human trafficking. Information Sciences. 2021;550:268–284. doi:https://doi.org/10.1016/j.ins.2020.10.029.

[43] Alvari H, Shakarian P, Snyder JK. A non-parametric learning approach to identify online human trafficking. In: 2016 IEEE Conference on Intelligence and Security Informatics (ISI). IEEE; 2016. p. 133–138.

[44] Alvari H, Shakarian P, Snyder JK. Semi-supervised learning for detecting human trafficking. Security Informatics. 2017;6(1):1–14. doi:https://doi.org/10.1186/s13388-017-0029-8.

[45] Andrews S, Brewster B, Day T. Organised crime and social media: A system for detecting, corroborating and visualising weak signals of organised crime online. Security Informatics. 2018;7(1):1–21. doi:https://doi.org/10.1186/s13388-018-0032-8.

[46] Ardakani HM. Identifying human trafficking networks in Louisiana by using authorship attribution and network modeling. Louisiana State University; 2020. Available from: https://digitalcommons.lsu.edu/gradschool_dissertations/5274/.

[47] Asif M, Kattan DA, Pamučar D, Ali G. q-Rung orthopair fuzzy matroids with application to human trafficking. Discrete Dynamics in Nature and Society. 2021;doi:https://doi.org/10.1155/2021/8261118.

[48] Binu M, Mathew S, Mordeson JN. Connectivity index of a fuzzy graph and its application to human trafficking. Fuzzy Sets and Systems. 2019;360:117–136. doi:https://doi.org/10.1016/j.fss.2018.06.007.

[49] Boyd DS, Perrat B, Li X, Jackson B, Landman T, Ling F, et al. Informing action for United Nations SDG target 8.7 and interdependent SDGs: Examining modern slavery from space. Humanities and Social Sciences Communications. 2021;8(1):1–14. doi:https://doi.org/10.1057/s41599-021-00792-z.

[50] Burbano D, Hernández-Alvarez M. Illicit, hidden advertisements on Twitter. In: 2018 International Conference on eDemocracy & eGovernment (ICEDEG). IEEE; 2018. p. 317–321.

[51] Caoli Jr A. Machine learning in the analysis of social problems: The case of global human trafficking. The British University in Dubai; 2019. Available from: https://bspace.buid.ac.ae/handle/1234/1579.

[52] Chambers N, Forman T, Griswold C, Lu K, Khastgir Y, Steckler S. Character-based models for adversarial phone extraction: Preventing human sex trafficking. In: Proceedings of the 5th Workshop on Noisy User-Generated Text (W-NUT 2019); 2019. p. 48–56.

[53] Chen Q, De Arteaga M, Herlands W. Canonical autocorrelation analysis and graphical modeling for human trafficking characterization; 2015. Available from: https://www.cs.cmu.edu/~epxing/Class/10715/project-reports/ChenDeArteagaHerlands.pdf.

[54] Cockbain E, Brayley H, Laycock G. Exploring internal child sex trafficking networks using social network analysis. Policing: A Journal of Policy and Practice. 2011;5(2):144–157. doi:https://doi.org/10.1093/police/par025.

[55] Cockbain E, Bowers K. Human trafficking for sex, labour and domestic servitude: How do key trafficking types compare and what are their predictors? Crime, Law and Social Change. 2019;72(1):9–34. doi:https://doi.org/10.1007/s10611-019-09836-7.

[56] Coxen J. A risk analysis and data driven approach to combating sex trafficking. University of Michigan; 2021. Available from: https://deepblue.lib.umich.edu/bitstream/handle/2027.42/170041/juliaoh_1.pdf.

[57] da Silva Santos M, Ladeira M, Van Erven GC, da Silva GL. Machine learning models to identify the risk of modern slavery in Brazilian cities. In: 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA). IEEE; 2019. p. 740–746.

[58] Darabian E, Borzooei RA. Results on vague graphs with applications in human trafficking. New Mathematics and Natural Computation. 2018;14(1):37–52. doi:https://doi.org/10.1142/S1793005718500047.

[59] de Vries I, Radford J. Identifying online risk markers of hard-to-observe crimes through semi-inductive triangulation: The case of human trafficking in the United States. The British Journal of Criminology. 2022;62(3):639–658. doi:https://doi.org/10.1093/bjc/azab077.

[60] Diaz M, Panangadan A. Natural language-based integration of online review datasets for identification of sex trafficking businesses. In: 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI). IEEE; 2020. p. 259–264.

[61] Diviák T, Dijkstra JK, van der Wijk F, Oosting I, Wolters G. Women trafficking networks: Structure and stages of women trafficking in five Dutch small-scale networks. European Journal of Criminology. 2021;doi:https://doi.org/10.1177/14773708211053135.

[62] Dubrawski A, Miller K, Barnes M, Boecking B, Kennedy E. Leveraging publicly available data to discern patterns of human-trafficking activity. Journal of Human Trafficking. 2015;1(1):65–85. doi:https://doi.org/10.1080/23322705.2015.1015342.

[63] Edge D, Yang W, Lytvynets K, Cook H, Galez-Davis C, Darnton H, et al. Design of a privacy-preserving data platform for collaboration against human trafficking. arXiv preprint arXiv:200505688. 2020;.

[64] Esfahani SS, Cafarella MJ, Pouyan MB, DeAngelo G, Eneva E, Fano AE. Context-specific language modeling for human trafficking detection from online advertisements. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics; 2019. p. 1180–1184.

[65] Gakiza J, Jilin Z, Chang Kc, Tao L. Human trafficking solution by deep learning with keras and OpenCV. In: International Conference on Advanced Intelligent Systems and Informatics. Springer; 2021. p. 70–79.

[66] Giacobe NA, Altmire JB, Forster AE, Jackson AC, Raibick EW, Reep JA, et al. Characterizing sex trafficking in Pennsylvania for law enforcement. In: 2016 IEEE Symposium on Technologies for Homeland Security (HST). IEEE; 2016. p. 1–5.

[67] Giommoni L, Ikwu R. Identifying human trafficking indicators in the UK online sex market. Trends in Organized Crime. 2021; p. 1–24. doi:https://doi.org/10.1007/s12117-021-09431-0.

[68] Goist M, Chen THY, Boylan C. Reconstructing and analyzing the transnational human trafficking network. In: 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). IEEE; 2019. p. 493–500.

[69] Granizo SL, Caraguay ÁLV, López LIB, Hernández-Álvarez M. Detection of possible illicit messages using natural language processing and computer vision on Twitter and linked websites. IEEE Access. 2020;8:44534–44546. doi:https://doi.org/10.1109/ACCESS.2020.2976530.

[70] Harmon R, Arnon D, Park B. TIP for tat: Political bias in human trafficking reporting. British Journal of Political Science. 2020; p. 1–11. doi:https://doi.org/10.1017/S0007123420000344.

[71] Hernández-Álvarez M. Detection of possible human trafficking in Twitter. In: 2019 International Conference on Information Systems and Software Technologies (ICI2ST). IEEE; 2019. p. 187–191.

[72] Hultgren M, Whitney J, Jennex ME, Elkins A. A knowledge management approach to identify victims of human sex trafficking. Communications of the Association for Information Systems. 2018;42(1):23. doi:https://doi.org/10.17705/1CAIS.04223.

[73] Ibanez M, Gazan R. Detecting sex trafficking circuits in the US through analysis of online escort advertisements. In: 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). IEEE; 2016. p. 892–895.

[74] Ibanez M, Suthers DD. Detecting covert sex trafficking networks in virtual markets. In: 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). IEEE; 2016. p. 876–879.

[75] Ibanez M, Suthers DD. Detection of domestic human trafficking indicators and movement trends using content available on open internet sources. In: 2014 47th Hawaii International Conference on System Sciences. IEEE; 2014. p. 1556–1565.

[76] Imperial JM. How do pedophiles tweet? Investigating the writing styles and online personas of child cybersex traffickers in the Philippines. arXiv preprint arXiv:210709881. 2021;.

[77] Kejriwal M, Ding J, Shao R, Kumar A, Szekely P. FlagIt: A system for minimally supervised human trafficking indicator mining. arXiv preprint arXiv:171203086. 2017;.

[78] Kejriwal M, Szekely P. Information extraction in illicit web domains. In: Proceedings of the 26th International Conference on World Wide Web; 2017. p. 997–1006.

[79] Kejriwal M, Szekely P. An investigative search engine for the human trafficking domain. In: International Semantic Web Conference. Springer; 2017. p. 247–262.

[80] Kejriwal M, Szekely P. Knowledge graphs for social good: An entity-centric search engine for the human trafficking domain. IEEE Transactions on Big Data. 2017;doi:https://doi.org/10.1109/TBDATA.2017.2763164.

[81] Kejriwal M, Szekely P, Knoblock C. Investigative knowledge discovery for combating illicit activities. IEEE Intelligent Systems. 2018;33(1):53–63. doi:https://doi.org/10.1109/MIS.2018.111144556.

[82] Kejriwal M, Kapoor R. Network-theoretic information extraction quality assessment in the human trafficking domain. Applied Network Science. 2019;4(1):1–26. doi:https://doi.org/10.1007/s41109-019-0154-z.

[83] Kejriwal M, Gu Y. Network-theoretic modeling of complex activity using UK online sex advertisements. Applied Network Science. 2020;5(1):1–23. doi:https://doi.org/10.1007/s41109-020-00275-1.

[84] Kennedy E. Predictive patterns of sex trafficking online. Carnegie Mellon University; 2012. Available from: https://kilthub.cmu.edu/articles/thesis/Predictive_Patterns_of_Sex_Trafficking_Online/6686309/1.

[85] Kiss L, Fotheringhame D, Mak J, McAlpine A, Zimmerman C. The use of Bayesian networks for realist evaluation of complex interventions: Evidence for prevention of human trafficking. Journal of Computational Social Science. 2020;4(1):25–48. doi:https://doi.org/10.1007/s42001-020-00067-8.

[86] Kulshrestha A. Detection of organized activity in online escort advertisements. McGill University (Canada); 2021. Available from: https://www.proquest.com/docview/2570357788?pq-origsite=gscholar&fromopenview=true.

[87] Lavelle-Hill R, Mazumder A, Goulding J, Smith G, Landman T. Machine learning methods for "Small-n, Large-p" problems: Understanding the complex drivers of modern-day slavery. Research Square preprint rs3rs-296275. 2021;doi:https://doi.org/10.21203/rs.3.rs-296275/v1.

[88] Lee MC, Vajiac C, Kulshrestha A, Levy S, Park N, Jones C, et al. InfoShield: Generalizable information-theoretic human-trafficking detection. In: 2021 IEEE 37th International Conference on Data Engineering (ICDE). IEEE; 2021. p. 1116–1127.

[89] Li L, Simek O, Lai A, Daggett M, Dagli CK, Jones C. Detection and characterization of human trafficking networks using unsupervised scalable text template matching. In: 2018 IEEE International Conference on Big Data (Big Data). IEEE; 2018. p. 3111–3120.

[90] Li R, Tobey M, Mayorga M, Caltagirone S, özaltın O. Detecting human trafficking: Automated classification of online customer reviews of massage businesses. Available at SSRN 3982796. 2021;.

[91] Libaque-Saenz CF, Lazo J, Lopez-Yucra KG, Bravo ER. Could machine learning improve the prediction of child labor in Peru? In: Lossio-Ventura JA, Alatrista-Salas H, editors. Information Management and Big Data. Springer International Publishing; 2018. p. 15–30.

[92] Liu Y, Zhu L, Szekely P, Galstyan A, Koutra D. Coupled clustering of time-series and networks. In: Proceedings of the 2019 SIAM International Conference on Data Mining. SIAM; 2019. p. 531–539.

[93] Malik DS, Mathew S, Mordeson JN. Fuzzy incidence graphs: Applications to human trafficking. Information Sciences. 2018;447:244–255. doi:https://doi.org/10.1016/j.ins.2018.03.022.

[94] Mancuso M. Not all madams have a central role: Analysis of a Nigerian sex trafficking network. Trends in Organized Crime. 2014;17(1):66–88. doi:https://doi.org/10.1007/s12117-013-9199-z.

[95] Mathew S, Mordeson JN. Fuzzy influence graphs. New Mathematics and Natural Computation. 2017;13(3):311–325. doi:https://doi.org/10.1142/S1793005717400129.

[96] Mathew S, Mordeson J, Yang HL. Incidence cuts and connectivity in fuzzy incidence graphs. Iranian Journal of Fuzzy Systems. 2019;16(2):31–43. doi:https://doi.org/10.22111/IJFS.2019.4540.

[97] McDonald GG, Costello C, Bone J, Cabral RB, Farabee V, Hochberg T, et al. Satellites can reveal global extent of forced labor in the world's fishing fleet. Proceedings of the National Academy of Sciences. 2021;118(3). doi:https://doi.org/10.1073/pnas.2016238117.

[98] Mensikova A, Mattmann CA. Ensemble sentiment analysis to identify human trafficking in web data. In: Proceedings of ACM Workshop on Graph Techniques for Adversarial Activity Analytics (GTA32018); 2018. p. 1–5.

[99] Mletzko D, Summers L, Arnio AN. Spatial patterns of urban sex trafficking. Journal of Criminal Justice. 2018;58:87–96. doi:https://doi.org/10.1016/j.jcrimjus.2018.07.008.

[100] Mordeson JN, Mathew S, Acharjee S. Aggregation operators: Applications to human trafficking and slavery. New Mathematics and Natural Computation. 2018;14(3):403–421. doi:https://doi.org/10.1142/S1793005718500242.

[101] Mordeson JN, Mathew S, Binu M. Dialectic synthesis: Application to human trafficking. New Mathematics and Natural Computation. 2019;15(3):395–410. doi:https://doi.org/10.1142/S1793005719500224.

[102] Mordeson JN, Mallenby M, Mathew S, Acharjee S. Human trafficking: Policy intervention. New Mathematics and Natural Computation. 2017;13(3):341–358. doi:https://doi.org/10.1142/S1793005717400142.

[103] Mordeson JN, Mathew S. T-Norm fuzzy graphs. New Mathematics and Natural Computation. 2018;14(1):129–143. doi:https://doi.org/10.1142/S1793005718500096.

[104] Mordeson JN, Mordeson JH, Mathew S. Directed graphs applied to human trafficking. New Mathematics and Natural Computation. 2018;14(3):445–455. doi:https://doi.org/10.1142/S1793005718500266.

[105] Mordeson JN, Schwab-McCoy A, Mathew S, Binu M. Fuzzy indices with applications to human trafficking. New Mathematics and Natural Computation. 2019;15(3):411–425. doi:https://doi.org/10.1142/S1793005719500236.

[106] Mordeson JN, Mathew S. Local look at human trafficking. New Mathematics and Natural Computation. 2017;13(3):327–340. doi:https://doi.org/10.1142/S1793005717400130.

[107] Mordeson JN, Mathew S. Non-deterministic flow in fuzzy networks and its application in identification of human trafficking chains. New Mathematics and Natural Computation. 2017;13(3):231–243. doi:https://doi.org/10.1142/S1793005717400087.

[108] Mordeson JN, Mathew S. Fuzzy logic applied to sustainable development goals and human trafficking. Symmetry. 2020;12(1):87. doi:https://doi.org/10.3390/sym12010087.

[109] Mordeson JN, Mathew S. Sustainable goals in combating human trafficking: Analysis by mathematics of uncertainty. Journal of Algebraic Hyperstructures and Logical Algebras. 2020;1(1):49–59. doi:https://doi.org/10.29252/HATEF.JAHLA.1.1.4.

[110] Mukherjee S, Sederholm T, Roman AC, Sankar R, Caltagirone S, Ferres JL. A machine learning pipeline for aiding school identification from child trafficking images. arXiv preprint arXiv:210605215. 2021;doi:https://doi.org/10.1145/3462203.3475924.

[111] Nagpal C, Miller K, Boecking B, Dubrawski A. An entity resolution approach to isolate instances of human trafficking online. arXiv preprint arXiv:150906659. 2015;.

[112] Poelmans J, Elzinga P, Ignatov DI, Kuznetsov SO. Semi-automated knowledge discovery: Identifying and profiling human trafficking. International Journal of General Systems. 2012;41(8):774–804. doi:https://doi.org/10.1080/03081079.2012.721662.

[113] Portnoff RS, Huang DY, Doerfler P, Afroz S, McCoy D. Backpage and bitcoin: Uncovering human traffickers. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2017. p. 1595–1604.

[114] Prashad L, Dutta M, Dash BM. Spatial analysis of child labour in India. Journal of Children's Services. 2021;16(4):269–280. doi:https://doi.org/10.1108/JCS-06-2019-0032.

[115] Price CN. Constraining international human trafficking: A network analysis. Creighton University; 2016. Available from: https://www.proquest.com/docview/1839269136?pq-origsite=gscholar&fromopenview=true.

[116] Progga FT, Shahria MT, Arisha A, Shanto MUA. A deep learning based approach to child labour detection. In: 2020 6th Information Technology International Seminar (ITIS). IEEE; 2020. p. 24–29.

[117] Rabbany R, Bayani D, Dubrawski A. Active search of connections for case building and combating human trafficking. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining; 2018. p. 2120–2129.

[118] Rodrigues DC, Prata DN, Silva MA. Exploring social data to understand child labor. International Journal of Social Science and Humanity. 2015;5(1):29. doi:https://doi.org/10.7763/IJSSH.2015.V5.416.

[119] Sabon L, Yang S, Zhang Q. Social network analysis of a latino sex trafficking enterprise. Journal of Human Trafficking. 2021; p. 1–17. doi:https://doi.org/10.1080/23322705.2021.1898830.

[120] Samanta S, Singhar SS, Gandomi AH, Ramasubbareddy S, Sankar S. A WiVi based IoT framework for detection of human trafficking victims kept in hideouts. In: International Conference on Internet of Things. Springer; 2020. p. 96–107.

[121] Sebastian A, Mordeson JN, Mathew S. Generalized fuzzy graph connectivity parameters with application to human trafficking. Mathematics. 2020;8(3):424. doi:https://doi.org/10.3390/math8030424.

[122] Sethi RJ, Gil Y, Jo H, Philpot A. Large-scale multimedia content analysis using scientific workflows. In: Proceedings of the 21st ACM International Conference on Multimedia; 2013. p. 813–822.

[123] Silva DR, Philpot A, Sundararajan A, Bryan NM, Hovy E. Data integration from open internet sources and network detection to combat underage sex trafficking. In: Proceedings of the 15th Annual International Conference on Digital Government Research; 2014. p. 86–90.

[124] Simonson E. Semi-supervised classification of social media posts: Identifying sex-industry posts to enable better support for those experiencing sex-trafficking. arXiv preprint arXiv:210403233. 2021;.

[125] Stylianou A, Xuan H, Shende M, Brandt J, Souvenir R, Pless R. Hotels-50k: A global hotel recognition dataset. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 33; 2019. p. 726–733.

[126] Szakonyi A, Chellasamy H, Vassilakos A, Dawson M. Using technologies to uncover patterns in human trafficking. In: ITNG 2021 18th International Conference on Information Technology-New Generations. Springer; 2021. p. 497–502.

[127] Szekely P, Knoblock CA, Slepicka J, Philpot A, Singh A, Yin C, et al. Building and using a knowledge graph to combat human trafficking. In: International Semantic Web Conference. Springer; 2015. p. 205–221.

[128] Tahir R, Imran MS, Minhas S, Sabahat N, Ilyas SHW, Gadi HR. Brick kiln detection and localization using deep learning techniques. In: 2021 International Conference on Artificial Intelligence (ICAI). IEEE; 2021. p. 37–43.

[129] Thöni A, Taudes A, Tjoa AM. An information system for assessing the likelihood of child labor in supplier locations leveraging Bayesian networks and text mining. Information Systems and e-Business Management. 2018;16(2):443–476. doi:https://doi.org/10.1007/s10257-018-0368-0.

[130] Tong E, Zadeh A, Jones C, Morency LP. Combating human trafficking with multimodal deep models. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Vancouver, Canada: Association for Computational Linguistics; 2017. p. 1547–1556.

[131] Upadhayay B, Lodhia ZAM, Behzadan V. Combating human trafficking via automatic OSINT collection, validation and fusion. In: Workshop Proceedings of the 15th International AAAI Conference on Web and Social Media; 2020.

[132] Vogt AA. Combating human trafficking using mathematics. In: Undergraduate Research and Scholarship Symposium. Duquesne University; 2016.Available from: mhttps://dsc.duq.edu/urss/2016/proceedings/3.

[133] Wang L, Laber E, Saanchi Y, Caltagirone S. Sex trafficking detection with ordinal regression neural networks. arXiv preprint arXiv:190805434. 2019;.

[134] Wang H, Cai C, Philpot A, Latonero M, Hovy EH, Metzler D. Data integration from open internet sources to combat sex trafficking of minors. In: Proceedings of the 13th Annual International Conference on Digital Government Research; 2012. p. 246–252.

[135] Whitney JC. KM vs human trafficking: An exploratory study on using emojis for a knowledge driven approach to identifying online human sex trafficking. San Diego State University; 2017. Available from: https://digitallibrary.sdsu.edu/islandora/object/sdsu%3A21609.

[136] White A, Guikema S, Carr B. Why are you here? Modeling illicit massage business location characteristics with machine learning. Journal of Human Trafficking. 2021; p. 1–21. doi:https://doi.org/10.1080/23322705.2021.1982238.

[137] Wiriyakun C, Kurutach W. Extracting co-occurrences of emojis and words as important features for human trafficking detection models. Journal of Intelligent Informatics and Smart Technology. 2021;7.

[138] Wiriyakun C, Kurutach W. Feature selection for human trafficking detection models. In: 2021 IEEE/ACIS 20th International Fall Conference on Computer and Information Science (ICIS Fall). IEEE; 2021. p. 131–135.

[139] Yang Y, Hu X, Liu H, Zhang J, Li Z, Yu PS. Understanding and monitoring human trafficking via social sensors: A sociological approach. arXiv preprint arXiv:180510617. 2018;.

[140] Yao Y, Liu Y, Guan Q, Hong Y, Wang R, Wang R, et al. Spatiotemporal distribution of human trafficking in China and predicting the locations of missing persons. Computers, Environment and Urban Systems. 2021;85:101567. doi:https://doi.org/10.1016/j.compenvurbsys.2020.101567.

[141] Zhou AJ, Luo J, McGibbney LJ. Multimedia metadata-based forensics in human trafficking web data. In: 2016 Workshop on Search and Exploration of X-rated Information. WSDM; 2016. p. 10–14.

[142] Zhu J, Li L, Jones C. Identification and detection of human trafficking using language models. In: 2019 European Intelligence and Security Informatics Conference (EISIC). IEEE; 2019. p. 24–31.

[143] Bhaumik A, Roy SK, Weber GW. Hesitant interval-valued intuitionistic fuzzy-linguistic term set approach in prisoners' dilemma game theory using TOPSIS: A case study on Human-trafficking. Central European Journal of Operations Research. 2020;28(2):797–816. doi:https://doi.org/10.1007/s10100-019-00638-9.

[144] Brelsford J, Parakh S. A systems modeling approach to analyzing human trafficking. In: 2018 Winter Simulation Conference (WSC). IEEE; 2018. p. 12–21.

[145] Dimas GL, Khalkhali ME, Bender A, Maass KL, Konrad R, Blom JS, et al. Estimating effectiveness of identifying human trafficking via data envelopment analysis. INFORMS Journal on Applied Analytics. 2023;doi:https://doi.org/10.1287/inte.2023.1162.

[146] Gerry QC F, Vallabhaneni S, Shaw P. Game theory and the human trafficking dilemma. Journal of Human Trafficking. 2021;7(2):168–186. doi:https://doi.org/10.1080/23322705.2019.1688086.

[147] Grimes J, Dillon RL, Tinsley CH. Systems dynamics as a method for analyzing human trafficking. 2011 Systems Dynamics Society. 2011;.

[148] Kapoor R, Kejriwal M, Szekely P. Using contexts and constraints for improved geotagging of human trafficking webpages. In: Proceedings of the Fourth International ACM Workshop on Managing and Mining Enriched Geo-Spatial Data; 2017. p. 1–6.

[149] Keskin BB, Bott GJ, Freeman NK. Cracking sex trafficking: Data analysis, pattern recognition, and path prediction. Production and Operations Management. 2021;30(4):1110–1135. doi:https://doi.org/10.1111/poms.13294.

[150] Konrad RA. Designing awareness campaigns to counter human trafficking: An analytic approach. Socio-Economic Planning Sciences. 2019;67:86–93. doi:https://doi.org/10.1016/j.seps.2018.10.005.

[151] Kosmas D, Sharkey TC, Mitchell JE, Maass KL, Martin L. Interdicting restructuring networks with applications in illicit trafficking. arXiv preprint arXiv:201107093. 2020;.

[152] Kosmas D, Melander C, Singerhouse E, Sharkey TC, Maass KL, Barrick K, et al. Generating synthetic but realistic human trafficking networks for modeling disruptions through transdisciplinary and community-based action research. arXiv preprint arXiv:220301893. 2022;.

[153] Kougkoulos I, Cakir MS, Kunz N, Boyd DS, Trautrims A, Hatzinikolaou K, et al. A multi-method approach to prioritize locations of labor exploitation for ground-based interventions. Production and Operations Management. 2021;doi:https://doi.org/10.1111/poms.13496.

[154] Kővári A, Pruyt E. A model-based exploration and policy analysis related to prostitution and human trafficking. International Journal of System Dynamics Applications (IJSDA). 2014;3(4):36–64. doi:https://doi.org/10.4018/ijsda.2014100103.

[155] Maass KL, Trapp AC, Konrad R. Optimizing placement of residential shelters for human trafficking survivors. Socio-Economic Planning Sciences. 2020;70:100730. doi:https://doi.org/10.1016/j.seps.2019.100730.

[156] Mahdiraji HA, Hafeez K, Jafarnejad A, Rezayar A. An analysis of the impact of negative CSR 'forced labour' parameter on the profitability of supply chain contracts. Journal of Cleaner Production. 2020;271:122274. doi:https://doi.org/10.1016/j.jclepro.2020.122274.

[157] Mahbub N. An approach to effective disruption of human trafficking through data-driven spatial-temporal prevalence estimation and optimal interdiction of interdependent illicit trades. State University of New York at Buffalo; 2021. Available from: https://www.proquest.com/docview/2582070644?pq-origsite=gscholar&fromopenview=true.

[158] Meier S, Vitor F. Developing a discrete event simulation model to overcome human trafficking. In: IIE Annual Conference. Proceedings. Institute of Industrial and Systems Engineers (IISE); 2021. p. 7–12.

[159] Neely B, Senft E, Turner B, Weeks B, Palmer J, Caddell J. Leveraging system dynamics and value modeling to identify strategic priorities to combat human trafficking in India. In: Proceedings of the Annual General Donald R. Keith Memorial Conference. Society for Industrial and Systems Engineering; 2019. p. 19–37.

[160] Ramchandani P, Bastani H, Wyatt E. Unmasking human trafficking risk in commercial sex supply chains with machine learning. Available at SSRN 3866259. 2021;doi:http://dx.doi.org/10.2139/ssrn.3866259.

[161] Senft E, Weeks B, Palmer J, Neely B, Turner B, Caddell J. A systems dynamics approach to human trafficking in Maharashtra, India. In: 2019 IEEE International Systems Conference (SysCon). IEEE; 2019. p. 1–7.

[162] Stapleton A, Chisholm N, Stapleton L. Sex trafficking from a supply chain systems perspective. IFAC Proceedings Volumes. 2012;45(10):15–20. doi:https://doi.org/10.3182/20120611-3-IE-4029.00006.

[163] Taylor KC. Teaching decision-making and building resilience in youth–A case study to reduce the supply of vulnerable youth to sex traffickers in Atlanta, Georgia. European Journal of Operational Research. 2018;268(3):960–970. doi:https://doi.org/10.1016/j.ejor.2017.11.067.

[164] Tezcan B, Maass KL. Human trafficking interdiction with decision dependent success. engrXiv preprint engrXiv:dt8fs. 2020;.

[165] Amin S. A step towards modeling and destabilizing human trafficking networks using machine learning methods. In: 2010 AAAI Spring Symposium Series; 2010.

[166] Borrelli D, Caltagirone S. Non-traditional cyber adversaries: Combatting human trafficking through data science. Cyber Security. 2020;4(1):77–90.

[167] Brewster B, Ingle T, Rankin G. Crawling open-source data for indicators of human trafficking. In: 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing. IEEE; 2014. p. 714–719.

[168] Brewster B, Polovina S, Rankin G, Andrews S. Knowledge management and human trafficking: Using conceptual knowledge representation, text analytics and open-source data to combat organized crime. In: International Conference on Conceptual Structures. Springer; 2014. p. 104–117.

[169] Deeb-Swihart J, Endert A, Bruckman A. Understanding law enforcement strategies and needs for combating human trafficking. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems; 2019. p. 1–14.

[170] Hundman K, Gowda T, Kejriwal M, Boecking B. Always lurking: Understanding and mitigating bias in online human trafficking detection. In: Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society; 2018. p. 137–143.

[171] Khanal S. Human trafficking in Nepal: Can big data help? Undergraduate Research Journal;24(1):5.

[172] LeBaron G. The role of supply chains in the global business of forced labour. Journal of Supply Chain Management. 2021;57(2):29–42.

[173] McKenzie J. Identifying and mitigating human trafficking risk through the use of financial institutions. Utica College; 2019. Available from: https://www.proquest.com/docview/2228305608?pq-origsite=gscholar&fromopenview=true.

[174] Orantes M. Leveraging machine learning and artificial intelligence to combat human trafficking. Utica College; 2018. Available from: https://www.proquest.com/docview/2160987666?pq-origsite=gscholar&fromopenview=true.

[175] Tambe P, Tambay P. Reducing modern slavery using AI and blockchain. In: 2020 IEEE/ITU International Conference on Artificial Intelligence for Good (AI4G). IEEE; 2020. p. 22–27.

[176] Weinberg N, Bora A, Sassetti F, Bryant K, Rootalu E, Bikziantieieva K, et al. AI against modern slavery: Digital insights into modern slavery reporting - challenges and opportunities. In: AI for Social Good, Association for the Advancement of Artificial Intelligence Fall Symposium 2020. AAAI; 2020.

[177] United States Justice Department. Justice Department leads effort to seize Backpage.com, the internet's leading forum for prostitution ads, and obtains 93-count federal indictment; 2018. Available from: https://www.justice.gov/opa/pr/justice-department-leads-effort-seize-backpagecom-internet-s-leading-forum-prostitution-ads.

[178] Sweileh WM. Research trends on human trafficking: A bibliometric analysis using Scopus database. Globalization and Health. 2018;14(1):1–12. doi:https://doi.org/10.1186/s12992-018-0427-9.

[179] Farrell A, Reichert J. Using US law-enforcement data: Promise and limits in measuring human trafficking. Journal of Human Trafficking. 2017;3(1):39–60. doi:https://doi.org/10.1080/23322705.2017.1280324.

[180] Gozdiak EM. Data and research on human trafficking: Bibliography of research-based literature. Washington, DC: Institute for the Study of International; 2011. Available from: https://www.ojp.gov/pdffiles1/nij/grants/224392.pdf.

[181] Laczko F. Human trafficking: The need for better data. Migration Information Source. 2002;1:61–80.

[182] Quintana NS, Rosenthal J, Krehely J. On the streets: The federal response to gay and transgender homeless youth. Washington, DC: Center for American Progress; 2010. Available from: https://www.americanprogress.org/article/on-the-streets/.

[183] Freeman L, Hamilton D. A count of homeless youth in New York City. New York: Empire State Coalition of Youth and Family Services; 2008. Available from: http://www.racismreview.com/downloads/HomelessYouth.pdf.

[184] Ali HM. Data collection on victims of human trafficking: An analysis of various sources. Journal of Human Security. 2010;6(1):55–69. doi:https://doi.org/10.3316%2Fjhs0601055.

[185] Aghazarm C, Laczko F. Human trafficking: New directions for research. Geneva International Organization for Migration; 2008. Available from: https://www.iom.int/sites/g/files/tmzbdl486/files/2018-07/human_trafficking_new_directions_for_research.pdf.

[186] Alvarez MB, Alessi EJ. Human trafficking is more than sex trafficking and prostitution: Implications for social work. Affilia. 2012;27(2):142–152. doi:https://doi.org/10.1177/0886109912443763.

[187] Gibbons P, Chisolm-Straker M, Stoklosa H. Human trafficking: Definitions, epidemiology, and shifting ground. Medical Perspectives on Human Trafficking in Adolescents. 2020; p. 1–12. doi:https://doi.org/10.1007/978-3-030-43367-3_1.

[188] Kerr P. Push and pull: The intersections of poverty, health disparities and human trafficking. Public Health & Social Justice. 2014;3(2):1–5.

[189] Gajic-Veljanoski O, Stewart DE. Women Trafficked Into Prostitution: Determinants, Human Rights and Health Needs. Transcultural Psychiatry. 2007;44(3):338–358. doi:https://doi.org/10.1177%2F1363461507081635.

[190] International Labour Office and Walk Free Foundation. Methodology of the global estimates of modern slavery: Forced labour and forced marriage; 2017. Available from: http://www.ilo.org/global/topics/forced-labour/%publications/WCMS_586127/langen/index.htm.

[191] U S Department of State. The 2019 trafficking in persons report; 2019. Available from: fromhttps://www.state.gov/reports/2019-trafficking-in-persons-report-2/nepal/.

[192] Feingold DA. Human trafficking. Foreign Policy; 2005. Available from: http://parlinfo.aph.gov.au/parlInfo/search/summary/summary.w3p;query=Id:%22library/jrnart/OD5H6%22.

[193] Logan TK, Walker R, Hunt G. Understanding Human Trafficking in the United States. Trauma, Violence, & Abuse. 2009;10(1):3–30. doi:https://doi.org/10.1177%2F1524838008327262.

[194] International Labour Office. Profits and poverty: The economics of forced labour; 2012. Available from: https://www.ilo.org/wcmsp5/groups/public/@ed_norm/@declaration/documents/publication/wcms_243027.pdf.

[195] Napolitano K. Human Trafficking: A Public Health Issue; 2016. Available from: https://combathumantrafficking.org/2016/12/human-trafficking-a-public-health-issue/.

[196] Greenbaum VJ. Commercial Sexual Exploitation and Sex Trafficking of Children in the United States. Current Problems in Pediatric and Adolescent Health Care. 2014;44(9):245–269. doi:https://doi.org/10.1016%2Fj.cppeds.2014.07.001.

[197] Denton E. International News Coverage of Human Trafficking Arrests and Prosecutions: A Content Analysis. Women & Criminal Justice. 2010;20(1-2):10–26. doi:https://doi.org/10.1080%2F08974451003641321.

[198] Foot K. In: Winterdyk J, Jones J, editors. Multisector Collaboration Against Human Trafficking. Cham: Springer International Publishing; 2020. p. 659–672.

[199] Cockayne J. Unshackling development: Why we need a global partnership to end modern slavery; 2015. Available from: ://freedomfund.org/wpcontent/uploads/Unshackling-Development-Final-For-Email.pdf.

[200] Moses S. The new abolitionism: Who's giving who to end modern day slavery?' inside philanthropy; 2016. Available from: https://www.insidephilanthropy.com/home/2016/3/3/funding-the-new-abolitionism-whos-giving-to-end-modern-day-s.html.

[201] Idris I. Funding of Anti-Modern Slavery Interventions. K4D Helpdesk Report. Brighton, UK: Institute of Development Studies; 2017. Available from: https://k4d.ids.ac.uk/resource/funding-of-anti-modern-slavery-interventions/.

[202] Foot KA, Toft A, Cesare N. Developments in Anti-Trafficking Efforts: 2008–2011. Journal of Human Trafficking. 2015;1(2):136–155. doi:https://doi.org/10.1080%2F23322705.2014.980663.

[203] Davy D. Anti–Human Trafficking Interventions. American Journal of Evaluation. 2016;37(4):486–504. doi:https://doi.org/10.1177%2F1098214016630615.

[204] Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Information Science and Systems. 2014;2(1). doi:https://doi.org/10.1186%2F2047-2501-2-3.

[205] Crowley ST, Meyer L. Sparking Innovation To Improve the Lives of People with Kidney Disease. Clinical Journal of the American Society of Nephrology. 2017;12(9):1548–1550. doi:https://doi.org/10.2215%2Fcjn.04420417.

[206] Mandinach EB. A Perfect Time for Data Use: Using Data-Driven Decision Making to Inform Practice. Educational Psychologist. 2012;47(2):71–85. doi:https://doi.org/10.1080%2F00461520.2012.667064.

[207] United Nations High Commissioner for Refugees. Global Trends: Forced Displacement; 2013. Available from: https://www.unhcr.org/us/media/unhcr-global-trends-2013.

[208] Boswell R, Schoderbek D, Collett TS, Ohtsuki S, White M, Anderson BJ. The Iġnik Sikumi Field Experiment, Alaska North Slope: Design, Operations, and Implications for CO2–CH4 Exchange in Gas Hydrate Reservoirs. Energy & Fuels. 2016;31(1):140–153. doi:https://doi.org/10.1021%2Facs.energyfuels.6b01909.

[209] Okun MA, Yeung EW, Brown S. Volunteering by older adults and risk of mortality: A meta-analysis. Psychology and Aging. 2013;28(2):564–577. doi:https://doi.org/10.1037%2Fa0031519.

[210] Jorm AF. Using the Delphi expert consensus method in mental health research. Australian & New Zealand Journal of Psychiatry. 2015;49(10):887–897. doi:https://doi.org/10.1177%2F0004867415600891.

[211] Berenguer G, Iyer AV, Yadav P. Disentangling the efficiency drivers in country-level global health programs: An empirical study. Journal of Operations Management. 2016;45(1):30–43. doi:https://doi.org/10.1016%2Fj.jom.2016.05.005.

[212] Zimmerman C, Hossain M, Watts C. Human trafficking and health: A conceptual model to inform policy, intervention and research. Social Science & Medicine. 2011;73(2):327–335. doi:https://doi.org/10.1016/j.socscimed.2011.05.028.

[213] Hudlow J. Fighting Human Trafficking Through Transit Monitoring: A Data-Driven Model Developed in Nepal. Journal of Human Trafficking. 2015;1(4):275–295. doi:https://doi.org/10.1080%2F23322705.2015.1098521.

[214] Joffres C, Mills E, Joffres M, Khanna T, Walia H, Grund D. Sexual slavery without borders: trafficking for commercial sexual exploitation in India. International Journal for Equity in Health. 2008;7(1). doi:https://doi.org/10.1186%2F1475-9276-7-22.

[215] United Nations Development Programme. Human Development Index; 2020. Available from: https://hdr.undp.org/data-center/country-insights#/ranks.

[216] Hennink M, Simkhada P. Sex Trafficking in Nepal: Context and Process. Asian and Pacific Migration Journal. 2004;13(3):305–338. doi:https://doi.org/10.1177%2F011719680401300302.

[217] Datta P. Nepali Female Migration and Trafficking. Journal of Social Sciences. 2005;11(1):49–56. doi:https://doi.org/10.1080%2F09718923.2005.11892493.

[218] Perry KM, McEwing L. How do social determinants affect human trafficking in Southeast Asia, and what can we do about it? A systematic review. Health & Hum Rts. 2013;15:138.

[219] Deane T. Cross-Border Trafficking in Nepal and India—Violating Women's Rights. Human Rights Review. 2010;11(4):491–513. doi:https://doi.org/10.1007%2Fs12142-010-0162-y.

[220] Kaufman MR, Crawford M. Research and Activism Review: Sex Trafficking in Nepal: A Review of Intervention and Prevention Programs. Violence Against Women. 2011;17(5):651–665. doi:https://doi.org/10.1177%2F1077801211407431.

[221] Geiger M, Pécoud A, editors. The Politics of International Migration Management. Palgrave Macmillan UK; 2010. Available from: https://doi.org/10.1057%2F9780230294882.

[222] Overseas Security Advisory Council. "Health Alert: Nepal, Lifting of Government Restrictions with Exceptions."; 2020. Available from: https://www.osac.gov/Country/Nepal/Content/Detail/Report/1ab67650-a24c-4831-abd6-193d57998817.

[223] Kansakar VBS. Nepal-India open borders: Prospects, problems and challenges. Series of Seminars, the Institute of Foreign Affairs and FES, Nepalgunj, Birgunj, Biratnagar and Kathmandu; 2001.

[224] Uddin MB. Human Trafficking in South Asia: Issues of Corruption and Human Security. International Journal of Social Work and Human Services Practice. 2014;2(1):18–27. doi:https://doi.org/10.13189%2Fijrh.2014.020103.

[225] Walk Free Foundation. The Global Slavery Index; 2018. Available from: https://downloads.globalslaveryindex.org/ephemeral/GSI-2018_FNL_190828_CO_DIGITAL_P-1640271480.pdf.

[226] Dubrawski A, Miller K, Barnes M, Boecking B, Kennedy E. Leveraging Publicly Available Data to Discern Patterns of Human-Trafficking Activity. Journal of Human Trafficking. 2015;1(1):65–85. doi:https://doi.org/10.1080%2F23322705.2015.1015342.

[227] Konrad RA, Trapp AC, Palmbach TM, Blom JS. Overcoming human trafficking via operations research and analytics: Opportunities for methods, models, and applications. European Journal of Operational Research. 2017;259(2):733–745. doi:https://doi.org/10.1016%2Fj.ejor.2016.10.049.

[228] Maass KL, Trapp AC, Konrad R. Optimizing placement of residential shelters for human trafficking survivors. Socio-Economic Planning Sciences. 2020;70:100730. doi:https://doi.org/10.1016%2Fj.seps.2019.100730.

[229] Ahn H, Afsharian M, Emrouznejad A, Banker R. Recent developments on the use of DEA in the public sector. Socio-Economic Planning Sciences. 2018;61:1–3. doi:https://doi.org/10.1016/j.seps.2017.06.001.

[230] Alda E, Cuesta J. Measuring the efficiency of humanitarian aid. Economics Letters. 2019;183:108618. doi:https://doi.org/10.1016/j.econlet.2019.108618.

[231] Izadikhah M, Azadi M, Kahi VS, Saen RF. Developing a new chance constrained NDEA model to measure the performance of humanitarian supply chains. International Journal of Production Research. 2019;57(3):662–682.

[232] Sun S. Measuring the relative efficiency of police precincts using data envelopment analysis. Socio-Economic Planning Sciences. 2002;36(1):51–71. doi:https://doi.org/10.1016%2Fs0038-0121%2801%2900010-6.

[233] Verma A, Gavirneni S. Measuring police efficiency in India: an application of data envelopment analysis. Policing: An International Journal of Police Strategies & Management. 2006;29(1):125–145. doi:https://doi.org/10.1108%2F13639510610648520.

[234] García-Sánchez IM. Evaluating the effectiveness of the Spanish police force through data envelopment analysis. European Journal of Law and Economics. 2007;23(1):43–57. doi:https://doi.org/10.1007%2Fs10657-007-9004-z.

[235] Verschelde M, Rogge N. An environment-adjusted evaluation of citizen satisfaction with local police effectiveness: Evidence from a conditional Data Envelopment Analysis approach. European Journal of Operational Research. 2012;223(1):214–225. doi:https://doi.org/10.1016%2Fj.ejor.2012.05.044.

[236] Vakkuri J. Research Techniques and Their Use in Managing Non-profit Organisations - An Illustration of DEA Analysis in NPO Environments. Financial Accountability and Management. 2003;19(3):243–263. doi:10.1111/1468-0408.00173.

[237] Kim H, Lee CW. Efficiency analysis for nonprofit organizations using DEA. Asia Pacific Journal of Innovation and Entrepreneurship. 2018;12(2):165–180. doi:https://doi.org/10.1108%2Fapjie-04-2018-0018.

[238] Özbek A. Efficiency Analysis of Non-Governmental Organizations Based in Turkey. International Business Research. 2015;8(9). doi:https://doi.org/10.5539%2Fibr.v8n9p95.

[239] Widiarto I, Emrouznejad A. Social and financial efficiency of Islamic microfinance institutions: A Data Envelopment Analysis application. Socio-Economic Planning Sciences. 2015;50:1–17. doi:https://doi.org/10.1016%2Fj.seps.2014.12.001.

[240] Nunamaker TR. Using data envelopment analysis to measure the efficiency of non-profit organizations: A critical evaluation. Managerial and Decision Economics. 1985;6(1):50–58. doi:https://doi.org/10.1002%2Fmde.4090060109.

[241] Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. European Journal of Operational Research. 1978;2(6):429–444. doi:https://doi.org/10.1016/0377-2217(78)90138-8.

[242] Lim S, Zhu J. DEA cross-efficiency evaluation under variable returns to scale. Journal of the Operational Research Society. 2015;66(3):476–487. doi:https://doi.org/10.1057%2Fjors.2014.13.

[243] Banker RD, Charnes A, Cooper WW. Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. Management Science. 1984;30(9):1078–1092. doi:https://doi.org/10.1287%2Fmnsc.30.9.1078.

[244] Podinovski VV, Thanassoulis E. Improving discrimination in data envelopment analysis: some practical suggestions. Journal of Productivity Analysis. 2007;28(1-2):117–126. doi:https://doi.org/10.1007%2Fs11123-007-0042-x.

[245] Cook WD, Tone K, Zhu J. Data envelopment analysis: Prior to choosing a model. Omega. 2014;44:1–4. doi:https://doi.org/10.1016/j.omega.2013.09.004.

[246] Medina-Borja A, Pasupathy KS, Triantis K. Large-scale data envelopment analysis (DEA) implementation: a strategic performance management approach. Journal of the Operational Research Society. 2007;58(8):1084–1098. doi:https://doi.org/10.1057%2Fpalgrave.jors.2602200.

[247] History com Editors. U.S. immigration timeline; 2018. Available from: https://www.history.com/topics/immigration/immigration-united-states-timeline.

[248] Davy M, Meyers DW, Batalova J. Who does what in U.S. immigration; 2005. Available from: https://www.migrationpolicy.org/article/who-does-what-us-immigration.

[249] National Immigration Forum. Fact Sheet: Immigration courts; 2018. Available from: https://immigrationforum.org/article/fact-sheet-immigration-courts/.

[250] TRAC Immigration. Immigration Court Backlog Tool: pending cases and length of wait in immigration courts;. Available from: https://trac.syr.edu/phptools/immigration/court_backlog/.

[251] American Immigration Council. An overview of U.S. refugee law and policy; 2021. Available from: https://www.americanimmigrationcouncil.org/research/overview-us-refugee-law-and-policy.

[252] Immigration and Nationality Act (INA), Title 8-Aliend and Nationality, 8 USC 1158: Asylum; 1965.

[253] Office of the Director. Guidance Regarding the Adjudication of Asylum Applications Consistent with INA § 208(d)(5)(A)(iii); 2018. Available from: https://www.justice.gov/eoir/page/file/1112581/download.

[254] The United States Department of Justice, Executive Office of Immigration Review. News and information; 2015. Available from: https://www.justice.gov/eoir/evolution-pre-1983.

[255] US Immigration com. U.S. Immigration Glossary; 2020. Available from: https://www.usimmigration.org/glossary.

[256] American Immigration Council. How the United States Immigration System Works; 2019. Available from: https://www.americanimmigrationcouncil.org/research/how-united-states-immigration-system-works.

[257] U S Citizenship and Immigration Services. Refugees and asylum; 2015. Available from: https://www.uscis.gov/humanitarian/refugees-and-asylum.

[258] United Nations High Commissioner for Refugees. Defensive asylum;. Available from: https://www.unhcr.org/en-us/defensive-asylum.html.

[259] American Immigration Council. Investigating the Immigration Court System; 2022. Available from: https://www.americanimmigrationcouncil.org/FOIA/investigating-immigration-court-system.

[260] Phillip K. Refugees, asylum seekers migrants: A crucial difference; 2018. Available from: https://www.habitatforhumanity.org.uk/blog/2016/09/refugees-asylum-seekers-migrants-crucial-difference/.

[261] Alvarez P. Trump Administration Puts Pressure on Completing Deportation Cases of Migrant Children; 2020. CNN Politics. Available from: https://www.cnn.com/2020/02/12/politics/unaccompanied-children-deportation/index.html.

[262] Nolo. What does my notice to appear (NTA) mean?; 2020. Available from: https://www.nolo.com/legal-encyclopedia/what-does-my-notice-appear-nta-mean.html.

[263] Executive Office for Immigration Review, Office of the Chief Immigration Judge. Immigration Court Practice Manual; 2016. Available from: https://www.justice.gov/eoir/page/file/1084851/download.

[264] Harchol-Balter M. Performance modeling and design of computer systems: queueing theory in action. Cambridge University Press; 2013. Available from: https://www.cs.cmu.edu/~harchol/PerformanceModeling/book.html.

[265] Stidham Jr S. Optimal design of queueing systems. Chapman and Hall/CRC; 2009.

[266] Eagly IV, Shafer S. A national study of access to counsel in immigration court. University of Pennsylvania Law Review. 2015;164(1):1–91.

[267] Ryo E. Predicting danger in immigration courts. Law & social inquiry. 2019;44(1):227–256.

[268] Koven SG, Götzke F. Theoretical Perspectives: Economics, Culture, Politics. In: American Immigration Policy. Springer; 2010. p. 19–33.

[269] Asad AL. Deportation decisions: Judicial decision-making in an American immigration court. American Behavioral Scientist. 2019;63(9):1221–1249.

[270] Rubio-Herrero J. The rise of unaccompanied alien children: Assessing the performance of a transition center with queuing theory. In: 2015 IEEE International Symposium on Technologies for Homeland Security (HST). IEEE; 2015. p. 1–6.

[271] Executive Office for Immigration Review. LOP cohort analysis; 2018. Available from: https://www.justice.gov/eoir/file/1091801/download.

[272] Siulc N, Byrne O, Son A, Cheng Z. Legal orientation program; 2008. Available from: https://www.vera.org/publications/legal-orientation-program-evaluation-and-performance-and-outcome-measurement-r

[273] The United States Department of Justice. FOIA Library;. Available from: https://www.justice.gov/eoir/foia-library-0.

[274] Ross SM. Introduction to probability models. Academic press; 2014.

[275] Lipton E, Preston J. As U.S. plugs border in Arizona, crossings shift to South Texas; 2013. Available from: https://www.nytimes.com/2013/06/17/us/as-us-plugs-border-in-arizona-crossings-shift-to-south-texas.html.

[276] TRAC Immigration. Pending cases and length of wait in Immigration Courts;. Available from: https://trac.syr.edu/phptools/immigration/court_backlog/.

[277] Harris LM. The one-year bar to asylum in the age of the immigration court backlog. Wis L REv. 2016; p. 1185.

[278] Garvin G. Immigration courts, lacking judges, are sinking under a massive backlog ...; 2017. Available from: https://www.miamiherald.com/news/local/immigration/article158889489.html.

[279] The White House, The United States Government. Remarks by president Biden in address to a joint session of Congress; 2021. Available from: http://www.whitehouse.gov/briefing-room/speeches-remarks/2021/04/29/remarks-by-president-biden-in-address-to-a-joint-session-of-congress/.

[280] TRAC Immigration. Immigration Court Backlog Now Growing Faster Than Ever, Burying Judges in an Avalanche of Cases;. Available from: https://trac.syr.edu/immigration/reports/675/.

[281] Hamrock E, Paige K, Parks J, Scheulen J, Levin S. Discrete event simulation for healthcare organizations: a tool for decision making. Journal of Healthcare Management. 2013;58(2):110–124.

[282] Günal MM, Pidd M. Discrete event simulation for performance modelling in health care: a review of the literature. Journal of Simulation. 2010;4(1):42–51. doi:https://doi.org/10.1057/jos.2009.25.

[283] Kampa A, Gołda G, Paprocka I. Discrete event simulation method as a tool for improvement of manufacturing systems. Computers. 2017;6(1):10. doi:https://doi.org/10.3390/computers6010010.

[284] Tako AA, Robinson S. The application of discrete event simulation and system dynamics in the logistics and supply chain context. Decision support systems. 2012;52(4):802–815. doi:https://doi.org/10.1016/j.dss.2011.11.015.

[285] de Azevedo Drummond P. Optimization of Port-of-Entry Operation in the US: An Anti-human Trafficking Focus. The University of Texas Rio Grande Valley; 2021. Available from: https://www.proquest.com/docview/2595490826?pq-origsite=gscholar&fromopenview=true.

[286] Irvine M, Coombs D, Skarha J, Del Pozo B, Rich J, Taxman F, et al. Modeling COVID-19 and its impacts on US Immigration and Customs Enforcement (ICE) detention facilities. Journal of Urban Health. 2020;97(4):439–447. doi:https://doi.org/10.1007/s11524-020-00441-x.

[287] Raman V, Vera C, Manna C. Bias, Consistency, and Partisanship in U.S. Asylum Cases: A Machine Learning Analysis of Extraneous Factors in Immigration Court Decisions. In: Equity and Access in Algorithms, Mechanisms, and Optimization. EAAMO '22. New York, NY, USA: Association for Computing Machinery; 2022.

[288] Ash E, Chen D, Andrus C, Godevais D, Ng G. Machine Prediction of Appeal Success in US Asylum Courts; 2020. Available from: https://users.nber.org/~dlchen/papers/Machine_Prediction_of_Appeal_Success_in_US_Asylum_Courts.pdf.

[289] Chen DL, Eagel J. Can machine learning help predict the outcome of asylum adjudications? In: Proceedings of the 16th edition of the International Conference on Articial Intelligence and Law; 2017. p. 237–240.

[290] Levesque C. Time is on their side? A duration analysis of US immigration court decisions. 2022;.

[291] Gandure J, Mhlanga S. Performance Analysis of Immigration Operation by discrete event modelling and simulation. Journal of Applied Science and Technology. 2009;14(1-2). doi:https://doi.org/10.4314/jast.v14i1-2.44344.

[292] Lim J, Azriani N, Nor M. Simulation by Queuing System at Immigration Department. International Academic Research Journal of Social Science. 2016;2(1):112–119.

[293] The United States Department of Justice, Executive Office of Immigration Review. News and information: Evolution of the U.S Immigration Court System: Pre-1983; 2015. Available from: https://www.justice.gov/eoir/evolution-pre-1983.

[294] TRAC Immigration. More Immigration Judges Leaving the Bench; 2020. Available from: https://trac.syr.edu/immigration/reports/617/.

[295] National Immigration Forum. Immigration Courts and Immigration Judges Fact Sheet; 2019. Available from: https://immigrationforum.org/wp-content/uploads/2019/04/Immigration-Courts-fact-sheet.pdf.

[296] United States Department of Justice, Office of Legislative and Public Affairs. The Executive Office for Immigration Review Swears in Six New Immigration Judges; 2010. Available from: https://www.justice.gov/sites/default/files/eoir/legacy/2010/09/20/IJInvestiture09172010.pdf.

[297] The United States, Office of Public Affairs. Executive Office for Immigration Review Announces Case Completion Numbers for Fiscal Year 2019; 2019. Available from: https://www.justice.gov/opa/pr/executive-office-immigration-review-announces-case-completion-numbers-fiscal-year-2019.

[298] STERN Law. How are immigration courts different than regular courts; 2017. Available from: https://sternlawfirm.us/blog/how-are-immigration-courts-different-than-regular-courts/.

[299] United States, Homeland Security. DHS and DOJ issue rule to efficiently and fairly process asylum claims; 2022. Available from: https://www.dhs.gov/news/2022/03/24/dhs-and-doj-issue-rule-efficiently-and-fairly-process-asylum-claims.

[300] Immigration Equality. Immigration basics; 2020. Available from: https://immigrationequality.org/asylum/immigration-basics/.

[301] The United States Department of Justice. Office of the chief immigration judge; 2022. Available from: https://www.justice.gov/eoir/office-of-the-chief-immigration-judge-bios#:~:text=The%20Office%20of%20the%20Chief,implementation%20for%20the%20immigration%20courts.

[302] Scherfke S, Lünsdorf O. Simpy Overview; 2020. Available from: https://simpy.readthedocs.io/en/latest/.

[303] United States Department of Justice, Executive Office for Immigration Review (EOIR). FY 2019 Budget Request At A Glance; 2019. Available from: https://www.justice.gov/jmd/page/file/1033216/download.

[304] Ramón C, O'Shea T. Why Hiring More Judges Would Reduce the Immigration Court Backlog; 2018. Available from: https://bipartisanpolicy.org/blog/why-hiring-more-judges-would-reduce-immigration-court-backlogs/.

[305] Rappaport N. No experience required: US hiring immigration judges who don't have any immigration law experience; 2020. Available from: https://thehill.com/opinion/immigration/481152-us-hiring-immigration-judges-who-dont-have-any-immigration-law-experience/.

[306] TRAC Immigration. A National Assessment of the Biden Administration's Dedicated Docket Initiative; 2022. Available from: https://trac.syr.edu/reports/704/.

[307] Communications and Legislative Affairs Division. Executive Office for Immigration Review Immigration Judge Training; 2022. Available from: https://www.justice.gov/eoir/page/file/1513996/download.

[308] Office of Public Affairs. DHS and DOJ announce dedicated docket process for more efficient immigration hearings; 2021. Available from: https://www.justice.gov/opa/pr/dhs-and-doj-announce-dedicated-docket-process-more-efficient-immigration-hearings.

[309] Srikantiah J. The immigration "rocket docket": Understanding the due process implications; 2014. Available from: https://law.stanford.edu/2014/08/15/the-immigration-rocket-docket-understanding-the-due-process-implications/.

[310] TRAC Immigration. With the Immigration Court's Rocket Docket Many Unrepresented Families Quickly Ordered Deported; 2018. Available from: https://trac.syr.edu/immigration/reports/441/.

[311] Vera Institute of Justice. Vera Institute of Justice Voices Opposition to "Rocket Docket" Court Proceedings; 2021. Available from: https://www.vera.org/newsroom/vera-institute-of-justice-voices-opposition-to-rocket-docket-court-proceedings.

[312] Reichlin-Melnick A. New Report Amplifies Concerns About 'Dedicated Docket' for Asylum Seekers; 2022. Available from: https://immigrationimpact.com/2022/01/20/dedicated-docket-asylum-seekers-problems/.

[313] Office of U S Government Accountability. Covid-19: Improvements needed in guidance and stakeholder engagement for Immigration Courts; 2021. Available from: https://www.gao.gov/products/gao-21-104404.

[314] Hernandez M. Immigration backlogs and Congressional funding; 2022. Available from: https://bipartisanpolicy.org/explainer/immigration-backlogs/#:~:text=Immigration%20Court%20Backlog&text=Though%20there%20are%20many%20factors,The%20administration%20deported.

[315] U S Citizenship and Immigration Services. Obtaining asylum in the United States; 2009. Available from: https://www.uscis.gov/humanitarian/refugees-and-asylum/asylum/obtaining-asylum-united-states.

Appendix A. Chapter 2 Estimating Effectiveness of Identifying Human Trafficking via Data Envelopment Analysis

Section A.1

Data Envelopment Analysis

Data envelopment analysis (DEA) dates to a seminal paper by [241] and provides a mechanism to compare similar entities, known in DEA as decision-making units (DMUs), to identify those that are exhibiting (in)efficiency relative to other DMUs. All DMUs are evaluated on their efficiency by their consumption of the same inputs to produce the same outputs in various quantities. The purpose of DEA is to allow for each DMU to express itself as efficient (with respect to comparable DMUs) through some combination of using its inputs to form outputs. Understanding the relative (in)efficiency of various DMUs enables best-practice recommendations to be developed for improving the operations of inefficient DMUs.

DEA models may consider returns to scale as constant (CRS) or variable (VRS). The former assumes that a change in the inputs results in a change of outputs in the same scale. The latter allows for nonconstant (that is, increasing or decreasing) returns to scale. In the current study, because there is no clear indication that a change in the inputs would result in the output change in the same scale, we consider the VRS model.

Suppose there are $|J|$ DMUs that consume $|M|$ inputs to produce $|N|$ outputs. Specifically, DMU $j \in J$ consumes x_{m_j} units of input $m \in M$ to produce y_{n_j} units of output $n \in N$. The output-oriented VRS DEA model maximizes the efficiency of DMU $j_0 \in J$ [243] by determining the optimal values of the input and output weights—v_m and u_n, respectively—in the following linear program (LP):

<u>Sets</u>

J	Set of decision-making units
M	Set of inputs
N	Set of outputs

<u>Parameters</u>

x_{m_j}	Amount of input $m \in M$ from unit $j \in J$
y_{n_j}	Amount of output $n \in N$ from unit $j \in J$

<u>Decision Variables</u>

w_m	Weight given to input $m \in M$
u_n	Weight given to output $n \in N$
v	Scale factor

<u>Output-Oriented Formulation</u>

$$\begin{aligned}
\text{Minimize} \quad & \sum_{m \in M} w_m x_{mj_0} - v \\
\text{Subject to:} \quad & \sum_{n \in N} u_n y_{nj_0} = 1, \\
& \sum_{m \in M} w_m x_{mj} - \sum_{n \in N} u_n y_{n_j} - v \geq 0, \quad \forall\, j \in J \\
& v \text{ free}, \\
& u_n, w_m \geq 0, \quad \forall\, n \in N,\ \forall\, m \in M.
\end{aligned}$$

For each DMU $j \in J$, denoted as the reference DMU j_0, the aforementioned LP is solved with respect to j_0. The goal is to determine weights, or price multipliers, that optimize the efficiency objective while ensuring that the total weighted input contributions for DMUs must be at least as large as the total weighted output contributions (so that DMUs cannot receive more in outputs than is put in).

If weights exist such that the weighted input contribution is equal to the weighted output contribution for the reference DMU (obtained through optimizing the LP), then this DMU is considered *efficient*. If this is not possible—that is, if weights that enable the DMU to get the value out that it puts in do not exist—then it is *inefficient*. For every inefficient DMU, there exists some combination of efficient DMUs that outperforms the inefficient DMU.

Section A.2

Sensitivity Analysis of Input and Output Features

We ran additional experiments across varying inputs and outputs to check the robustness of our final selected model. Table BB.1 details the three inputs and five outputs considered throughout our modeling process. We considered multiple combinations before settling on our final model, which is Model A in Table BB.2. The most salient experiments and results are presented in Tables BB.2 and BB.3, respectively.

Table BB.1: **Three Inputs and Five Outputs Were Considered in the DEA Model.**

Feature		Description
Inputs	Number of staff	The average number of staff working at a station per quarter
	Staff test scores	The average test scores of staff working at a station per quarter
	Hours worked by staff	The average weekly hours worked per staff in a quarter
Outputs	Count of IRFs	The number of IRF forms collected per quarter
	Count of VIFs	The number of VIF forms collected per quarter
	IRF completeness	The average completeness of IRF required questions per quarter
	VIF completeness	The average completeness of VIF required questions per quarter
	Percent of "correct" instances of trafficking	The percent of total VIF forms that are estimated to be a positive instance of trafficking *It is important to note this feature is limited by the data available and likely an overestimate.*

The first set of experiments, which we call *varying input* experiments, are those in which we varied only the inputs, holding constant the four outputs used in our final model. We note that transit monitoring (application of the IRF and VIF forms) has two objectives: (1) to interact with a multitude of individuals (the count of forms) and (2) to sufficiently complete the forms (completeness measure) to help determine those who are at higher risk of being trafficked. Even so, we carried out experiments to exclude both VIF and IRF form counts and only include the form completeness measures. These results did have an effect on the cross-efficiency rankings of stations (see Table BB.3). Moreover, five DMUs that are "efficient" under our final proposed model were not "efficient" in this experiment. LJI's operations are currently designed around the assumption that the more effective an interaction with individuals is (measuring in both counts and completeness), the better the chances of identifying a potential instance of trafficking and therefore aiding in the fight against human trafficking. Therefore, because we are only comparing stations with a similar amount of flow (people crossing the station), the count of forms is an important output measure for LJI's transit-monitoring efforts and is included in our final model. It should also be noted that although both Model A and E produce the same rankings, staff hours were deemed an important feature to include in this context, and therefore, the set of inputs considered hereafter are number of staff, staff test scores, and staff hours.

Appendix A. Chapter 2 Estimating Effectiveness of Identifying Human Trafficking via Data Envelopment Analysis

To test the sensitivity of this Model A with respect to the outputs, we ran *varying output* experiments that varied only the outputs, holding the inputs constant. One such output considered was an approximate estimate of the proportion of "correct instances" of trafficking. Given the limitations of available data that was collected, this value better represents an instance of an individual most likely to have been trafficked. This approximation was calculated using two features in the data, and although our estimate is far from perfect, this combined feature provides more information on those who are more likely to be a "correct instance" of trafficking. Using these data, we created an output feature that was the proportion of "correct instances" out of the total number of VIF forms filled out. The average percentage across all DMUs was just under 60%. After using this feature for robustness checking, we determined that the cross-efficiency rankings were consistent whether this feature was included or excluded. Because the results did not change and this feature was an approximation, we retained the simpler Model A (excluding this new output feature).

Table BB.2 describes each model considered. We compared these different models on the cross-efficiency scores of our final output-oriented VRS model (base model). The results are presented in Table BB.3; the parenthetical number is the ranking for the highest cross-efficiency, with 1 indicating the highest (most) station efficiency, and 7 indicating the lowest (least). Bolded values indicate instances in which a station's ranking was impacted by model variation. These experiments resulted in Model A being the best-suited model for LJI and thus the final model presented in our study.

Table BB.2: **Twelve Experiments Were Considered by Varying the Input and Outputs.**

	Model	Input values			Output values				
		Number of staff	Staff test scores	Staff hours	IRF forms	VIF forms	IRF completion	VIF completion	"Correct" instance of trafficking
	A	•	•	•	•	•	•	•	
Varying inputs	B	•			•	•	•	•	
	C		•		•	•	•	•	
	D			•	•	•	•	•	
	E	•	•		•	•	•	•	
	F	•		•	•	•	•	•	
	G		•	•	•	•	•	•	
Varying outputs	H	•	•	•			•	•	
	I	•	•	•			•	•	•
	J	•	•	•	•	•	•	•	•
	K	•	•	•	•	•			•
	L	•	•	•	•	•			

Table BB.3: **The Cross-Efficiency of the 12 Experiments, Revealing which Stations Have the Highest and Lowest Cross-Efficiency.**

		Station						
	Model	Nepalgunj	Mahendranagar	Karkarvitta	Birgunj	Biratnagar	Bhairawa	Bhadrapur
	A	**0.885 (6)**	**0.929 (2)**	**0.943 (1)**	**0.916 (5)**	**0.922 (4)**	**0.925 (3)**	**0.856 (7)**
Varying inputs	B	**0.927 (5)**	**0.922 (6)**	**0.949 (3)**	**0.949 (2)**	**0.950 (1)**	**0.944 (4)**	0.901 (7)
Varying inputs	C	0.891 (6)	**0.932 (4)**	0.953 (1)	0.924 (5)	**0.935 (2)**	0.934 (3)	0.868 (7)
Varying inputs	D	0.913 (6)	**0.914 (5)**	0.954 (1)	**0.919 (4)**	**0.949 (2)**	0.939 (3)	0.883 (7)
Varying inputs	E	0.893 (6)	0.933 (2)	0.945 (1)	0.929 (5)	0.929 (4)	0.931 (3)	0.867 (7)
Varying inputs	F	0.920 (6)	**0.924 (5)**	0.948 (1)	**0.938 (4)**	**0.946 (2)**	0.939 (3)	0.891 (7)
Varying inputs	G	0.881 (6)	**0.926 (4)**	0.949 (1)	0.908 (5)	**0.927 (3)**	**0.928 (2)**	0.854 (7)
Varying outputs	H	0.917 (6)	**0.947 (3)**	0.956 (1)	**0.947 (4)**	**0.952 (2)**	**0.944 (5)**	0.891 (7)
Varying outputs	I	0.891 (6)	0.936 (2)	0.937 (1)	**0.923 (4)**	**0.927 (3)**	**0.921 (5)**	0.861 (7)
Varying outputs	J	0.853 (6)	0.916 (2)	0.922 (1)	0.889 (5)	0.892 (4)	0.900 (3)	0.823 (7)
Varying outputs	K	0.240 (6)	0.433 (2)	0.528 (1)	0.301 (5)	0.304 (4)	0.431 (3)	0.823 (7)
Varying outputs	L	0.240 (6)	**0.418 (3)**	0.516 (1)	0.271 (5)	0.313 (4)	**0.442 (2)**	0.218 (7)

Appendix B. Chapter 4: Modeling the United States Immigration Court Using Discrete Event Simulation

Section B.1 Additional Figures

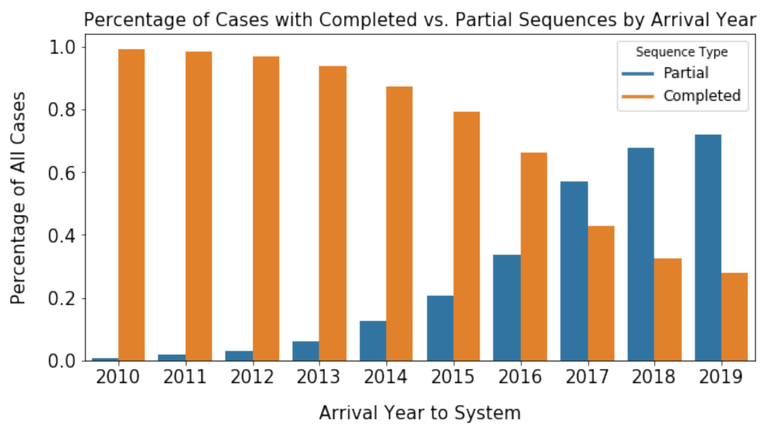

Figure 17: **Percentage of Cases with Completed Verses Partial Sequences by Arrival Year 2010-2019.**

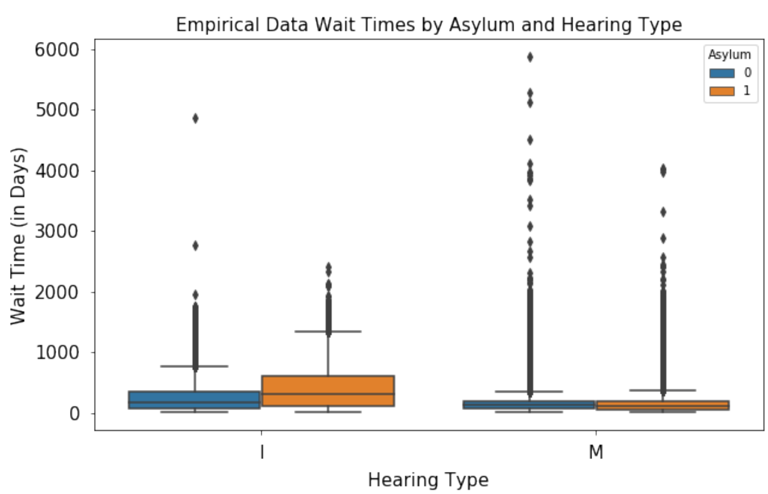

Figure 18: **Empirical Wait times for Asylum and Non-asylum cases by Hearing Type.**

Appendix C. Chapter 5: Assessing Policies to Improve the United States Immigration Court Operations Using Discrete Event Simulation

Section C.1 Policy 1: Varying the Number of Judges

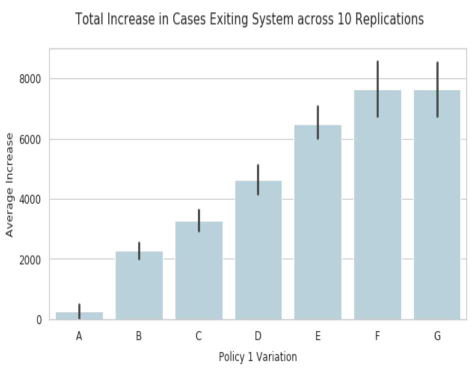

Figure 19: **Total Increase in Cases Exiting System across Policy 1 Variations.**

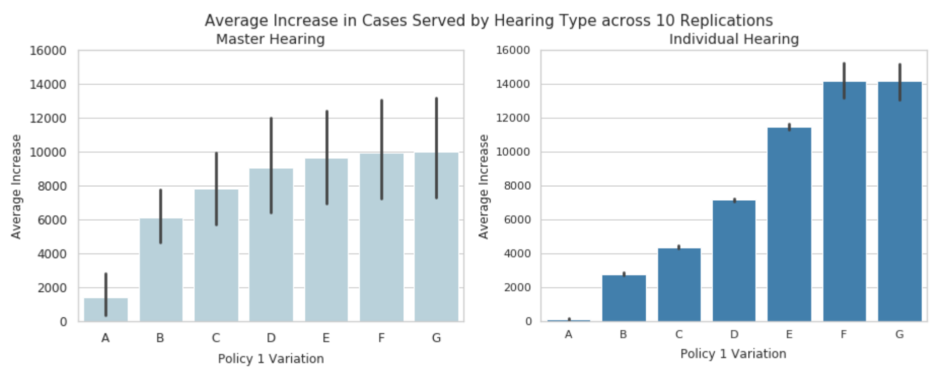

Figure 20: **Total Increase in Cases Each Hearing Type across Policy 1 Variations.**

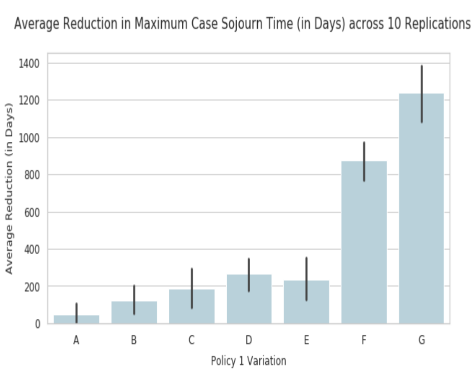

Figure 21: **Average Reduction in Maximum Sojourn Time across Policy 1 Variations.**

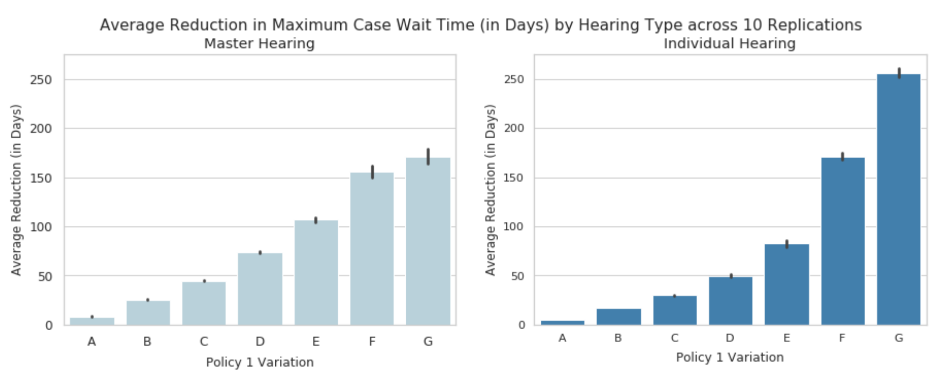

Figure 22: **Average Reduction in Maximum Case Wait Time for Each Hearing Type across Policy 1 Variations.**

Section C.2

Policy 2: Dedicated Dockets

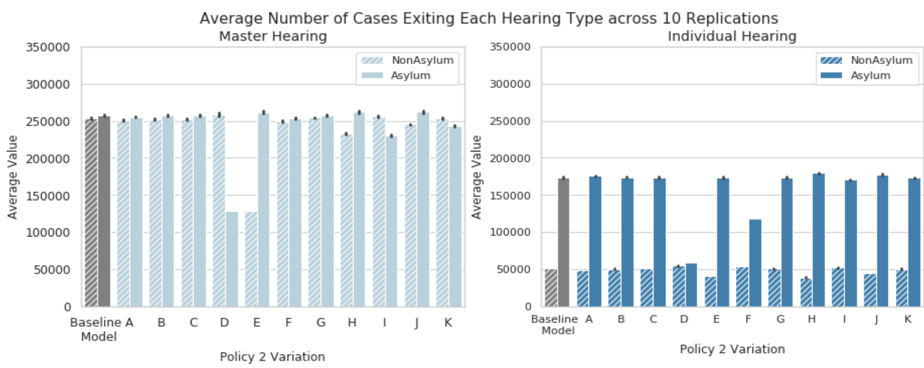

Figure 23: **Average Number of Cases Exiting each Hearing Type across Policy 2 Variations.**

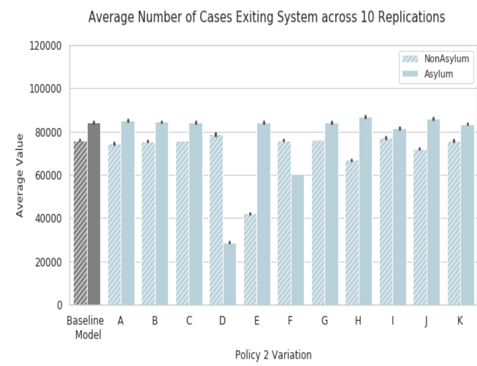

Figure 24: **Average Number of Cases Exiting the System across Policy 2 Variations.**

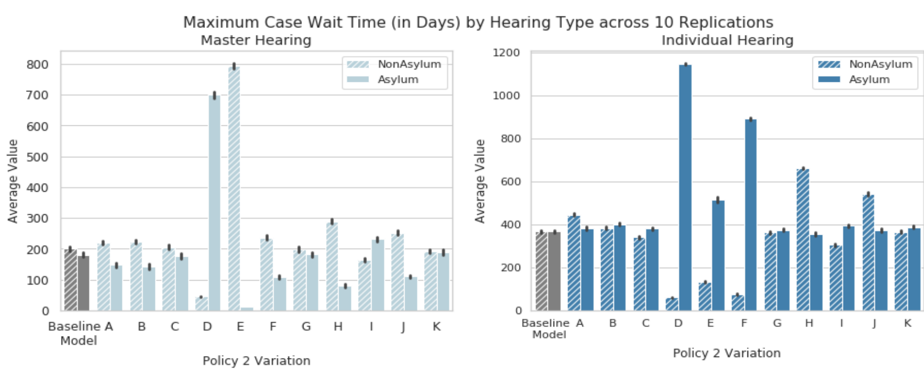

Figure 25: **Average Maximum Case Wait Time for Each Hearing Type across Policy 2 Variations.**

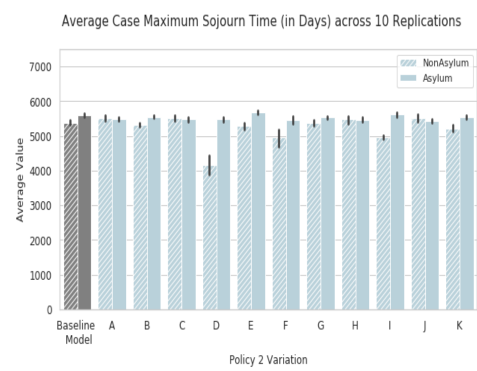

Figure 26: **Average Maximum Case Sojourn Time across Policy 2 Variations.**

Section C.3

Policy 3: Make-up Capacity

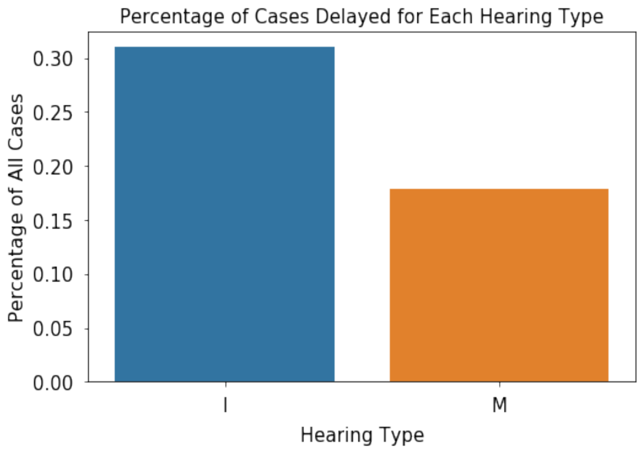

Figure 27: **Percentage of Hearings with a Delay for Each Hearing Type.**

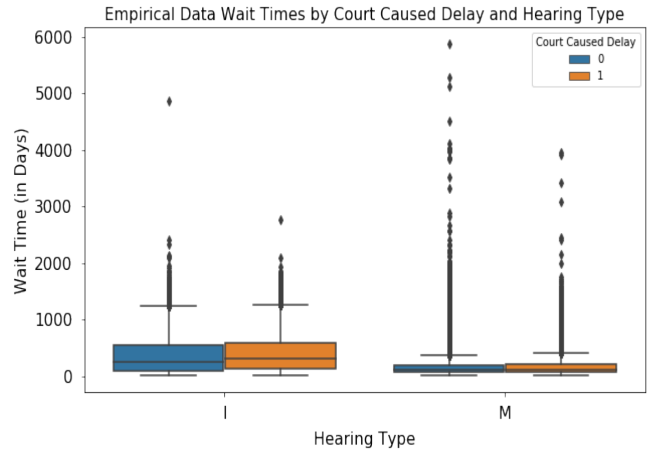

Figure 28: **Empirical Wait times by Court Caused Delays and Hearing Types.**

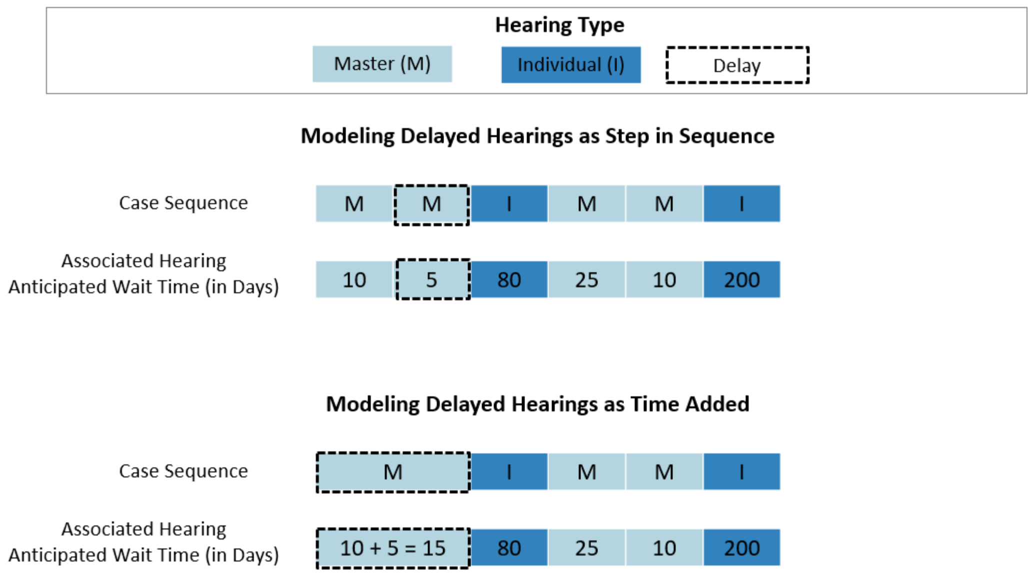

Figure 29: **Modeling Delayed Cases as Steps in a Sequence Verses as Time Added.**

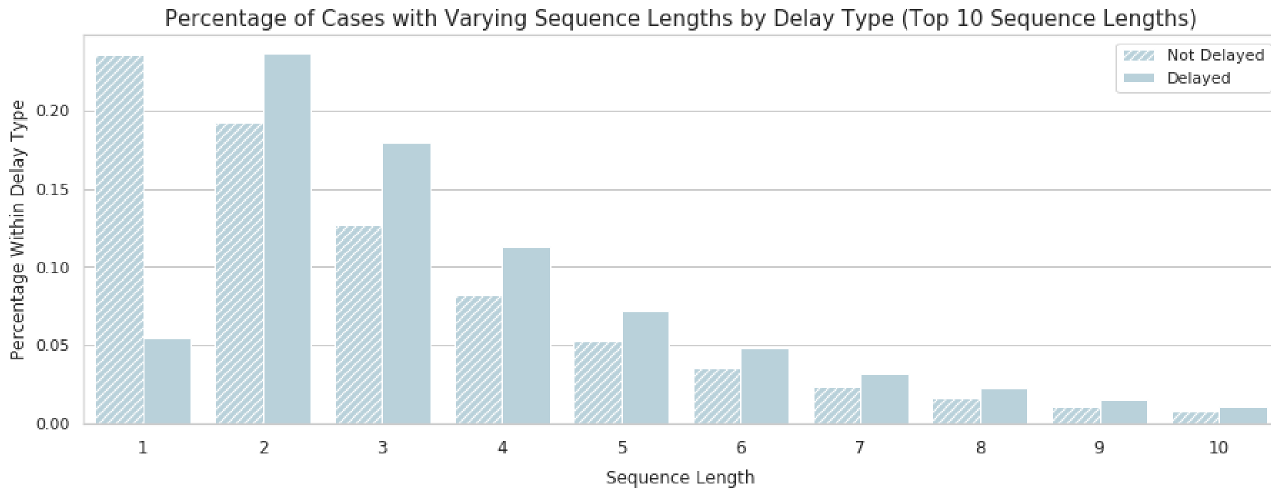

Figure 30: **Percentage of Cases with Varying Sequence Lengths by Delay Type (Top 10 Sequence Lengths).**

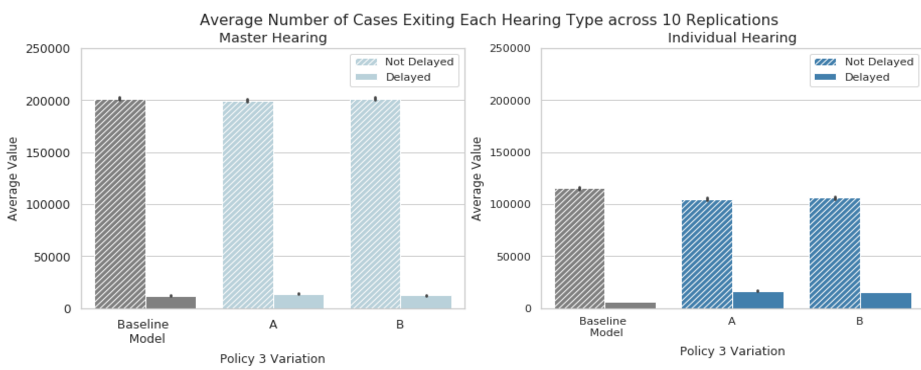

Figure 31: **Average Number of Cases Exiting Each Hearing across Policy 3 Variations.**

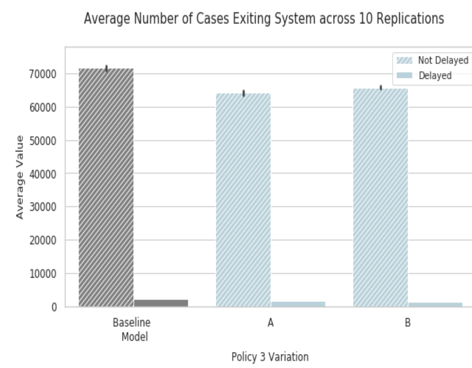

Figure 32: **Average Number of Cases Exiting System across Policy 3 Variations.**

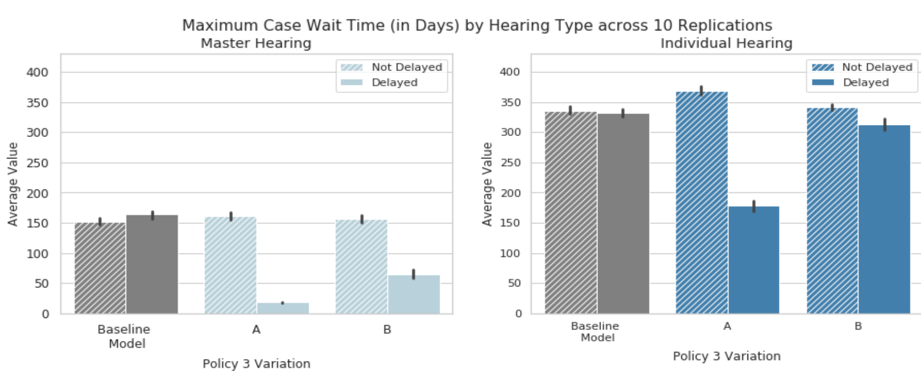

Figure 33: **Average Maximum Case Wait Time by Hearing Type across Policy 3 Variations.**

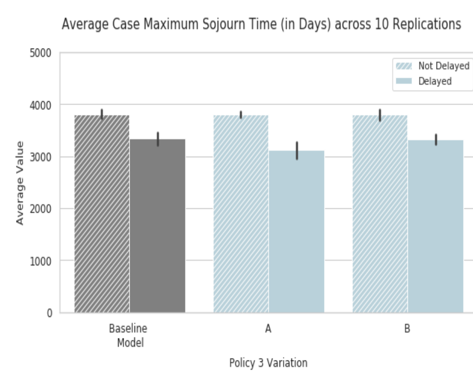

Figure 34: **Average Maximum Case Sojourn Time by Hearing Type across Policy 3 Variations.**