# PERSONALITIES OF DISASTER MANAGEMENT: DATA-DRIVEN APPROACHES TO QUANTIFYING RESILIENCE AND BEHAVIORAL UNCERTAINTY IN RESPONSE TO NATURAL HAZARDS

By

Paul M. Johnson

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY
in

Environmental Engineering

August 31, 2021

Nashville, Tennessee

Approved:

Mark Abkowitz, Ph.D.

Hiba Baroud, Ph.D.

Ralf Bennartz, Ph.D.

Corey Brady, Ph.D.

Janey Camp, Ph.D.

Jonathan Gilligan, Ph.D.

Craig Philip, Ph.D.

Dedicated to family, friends, and Shadow ... the smartest pup I’ve ever known.

## ACKNOWLEDGMENTS

I would like to especially thank my advisors, Dr. Mark Akowitz and Dr. Hiba Baroud, for their constant support and guidance throughout my time at Vanderbilt. Additionally, I would like to express my sincere gratitude to all members of my committee: Dr. Ralf Bennartz, Dr. Corey Brady, Dr. Janey Camp, Dr. Jonathan Gilligan, and Dr. Craig Philip. From each of you, I learned valuable skill sets that have made all of this work possible and that I will continue to use throughout the rest of my career. Financial support for this work was provided by the U.S. Department of Housing and Urban Development (HUD) Mississippi River Supply Chain Management project and by the IBM PhD Fellowship Program.

## TABLE OF CONTENTS

Page

## LIST OF TABLES

## LIST OF FIGURES

CHAPTER 1

# Introduction

Throughout human history, natural disasters have been synonymous with catastrophes (Fjord and Manderson, 2009). However, the term "natural" disaster is a misnomer because "every natural disaster involves a unique pattern of physical energy expenditure and human reaction" (Alexander, 2002, p. 289). People play a significant role in determining the outcomes of these events, thus illustrating the concept of community resilience (Fjord and Manderson, 2009; Johnson et al., 2020). Resilience is an important concept in the context of natural disasters because it helps reconcile why hazards of similar type and magnitude can have such disparate outcomes on different communities (Birkmann et al., 2014). In general, communities with greater resilience have a better chance of withstanding these events (Seaberg et al., 2017). While this sentiment also applies to man-made disasters, the focus of this dissertation is on how aspects of community resilience affect risks associated with natural hazards.

## 1.1 Disaster Risk and Community Resilience

The risk management literature defines *risk* as the interaction of *hazard* and *resilience* (or *vulnerability*) (Alexander, 2002; Birkmann, 2007). Figure 1.1 illustrates this concept. The term hazard pertains to characteristics of a threat itself (e.g. type, severity, and likelihood), and in the context of natural disasters, hazard can be viewed as the "natural" part of this coupled human-natural system (Birkmann, 2007). The other part of the risk equation, which is the focus of this dissertation, includes resilience and vulnerability.


Figure 1.1: The Risk Equation

Although definitions for resilience and vulnerability can vary across disciplines, in general, resilience describes a system's ability to mitigate, resist, and/or recover from harm, and vulnerability refers to a system's susceptibility to harm (Birkmann et al., 2014; Cutter et al., 2014; UNDP, 2011; UN/ISDR, 2005). There is still debate regarding the exact relationship between resilience and vulnerability. Some view resilience as simply the antithesis to vulnerability, but most recognize differences in the active vs. passive connotations embedded in the former and latter terms respectively (Cutter et al., 2008, 2014; Mohammad and El-adaway, 2017; Norris et al., 2008; Pendall et al., 2010). However, as Bakkensen et al. (2017) stated, from a practical standpoint, both aim to achieve the same goal in helping communities reduce the impacts of disasters. We adopt a similar mindset throughout this dissertation, recognizing that the two concepts may differ subtly in theory but are fundamentally the same in their practical application of risk management. As such, unless explicitly stated otherwise, readers can infer that the term "community resilience" also incorporates "community vulnerability" (i.e., aspects of a community that characterize an increased resilience to natural disasters also imply a decreased vulnerability to these events).

## 1.2 Motivation of Study

Thanks to modern technology, society's ability to understand and anticipate natural hazards has greatly improved over the last several decades, albeit to degrees of aleatory and epistemic uncertainty. However, in 2005 at the World Conference on Disaster Reduction, it became clear that research on community resilience was lacking comparatively, thus spurring efforts to better investigate this part of the risk equation (UN/ISDR, 2005). Here, use data-driven approaches to improve assessments of community resilience.

In particular, much work has gone into trying to empirically define community resilience with little proven success. Community resilience is multidimensional and can be contextually dependent on the type of disaster (Sharifi, 2016; Sorg et al., 2018). Furthermore, it is spatially dependent with respect to both location and scale and varies temporally, even within the time-frame of a single disaster (Adger, 2006; Nelson et al., 2015; Ritchie and Gill, 2007). Consequently, there has been a proliferation of frameworks on how to best measure community resilience, but the vast majority lack any construct validity (i.e., demarcations of the degree to which a model actually measures what it claims to be measuring) (Bakkensen et al., 2017). As such, functionally comparing approaches is impractical because there is a lack of objective justification for the use of one framework over another.

Similarly, models that attempt to simulate the impacts of natural disasters on communities rarely have practical use because they fail to incorporate how human decision-making in preparation and/or response to such events affects community resilience or lack empirical justification for assumptions thereof (Fiedrich and Burghardt, 2007; Janssen and Ostrom, 2006; Taylor et al., 2014). In general, calibrating and validating models pertaining to disaster risk management is difficult because observational data is lacking for extreme weather events (He and Liu, 2012; Jackson et al., 2016). Cognitive and behavioral models informed by decision theory or psychology can provide solid rationale for decision-making paradigms in simulations, but it can be difficult to ensure that the simulated behavior accurately reflects those represented by the various theories and that the decision models are appropriate for the given application (Abdulkareem et al., 2018; Jumadi et al., 2018; Stefanelli and Seidl, 2017; Ahlqvist et al., 2018; Massey et al., 2018). This problem extends beyond the field of risk management, and in recent years, there as been an increasing drive among researchers to devise better approaches for empirically grounding realistic decision-making processes in simulations (Janssen and Ostrom, 2006; Abdulkareem et al., 2018; Choi and Lee, 2018).

### 1.3 Overview of Study

The work presented in this dissertation helps address both of these knowledge gaps. First, we develop a data-driven approach for reconciling a popular class of frameworks aimed at measuring community resilience known as resilience indices. Our methodology establishes an objective and intuitive schema for relating the constituent elements of well-known resilience indices, which in turn helps researchers establish construct validity in the field. Second, we use advanced machine techniques to develop empirically-based decision models that can be encoded in simulations to better represent how decision-makers respond to the threat of natural hazards and to more accurately quantify how their behaviors affect community outcomes. We present insights drawn from our decision model in the context of previous findings from psychometrics and decision theory. Additionally, we develop a use case by applying our decision model to an agent-based model (ABM) that simulates economic production losses suffered as a consequence of decisions made in response to inland waterway disruptions due to floods.

This dissertation is organized as follows. Chapter 2 describes our data-driven approach for reconciling community resilience indices. Chapter 3 presents a behavioral study we conducted where community resilience is conceptualized as a common-pool resource (CPR) game and discusses how we trained Bayesian Additive Regression Tree (BART) models on results to predict participants' responses. Chapter 4 describes an empirically-based ABM we developed that simulates how businesses along the Mississippi River may reasonably reroute inland waterway shipments in response to various flood scenarios and calculates the subsequent economic impacts of such decisions. Chapter 5 discusses an ABM we developed that combines findings from Chapters 3 and 4 to simulate outcomes of a CPR dilemma that arises when state decision-makers choose to invest in the development of a publicly operated, flood-resilient port that can lead to savings in economic production losses suffered during inland waterway disruptions due to floods. Chapter 6 contains our concluding remarks and directions for future research. It should be noted that Chapters 2-5 consist of published or soon to be published journal articles that are intended to be read as separate documents, so they contain slight overlaps in material.

# CHAPTER 2

## A factor analysis approach toward reconciling community resilience indices for natural hazards

The concept of resilience helps explain why natural hazards of similar type and magnitude can have disparate impacts on varying communities. Numerous frameworks have been developed to measure community resilience, but a clear and consistent method of comparing them is lacking. Here, we develop a data-driven approach for reconciling a popular class of frameworks known as resilience indices. In particular, we conduct an exploratory factor analysis on a comprehensive set of variables from established indices measuring community resilience at the US county level. The resulting factor model suggests that 50 of the 130 analyzed variables effectively load onto five dimensions: wealth, poverty, agencies per capita, elderly populations, and non-English speaking populations. These factors establish an objective and intuitive schema for relating the constituent elements of resilience indices, thus providing researchers a flexible yet robust framework for validating and expanding upon current approaches.

### 2.1 Introduction

While efforts to thematically define community resilience have been fairly successful, attempts to measure this concept are still nascent. Some researchers surmise that community resilience does not lend itself to being measured (Ayyub, 2015; Park et al., 2013). Others suggest otherwise, and a popular approach they use to quantify community resilience entails the creation of an *index*. An index is a geographical composition of *indicators*, delineations of what should theoretically constitute community resilience. Examples of indicators include constructs such as transportation capacity, medical care capacity, and gender income equality. In turn, these indicators are measured with corresponding data, such as percent households with at least one vehicle, hospital beds per 10,000 people, and difference between male and female median incomes (Cutter et al., 2014).

The authority of indicators is largely predicated on the notion of capital systems, which are theory-driven categorizations of community resources such as social, institutional, economic, and environmental assets (Norris et al., 2008; Rowcliffe et al., 1999; Twigg, 2007). Stemming from the postulate that those with fewer resources are disproportionately afflicted by catastrophic events, indicators are constructed to help inform what aspects of capital systems a community might be lacking, thereby decreasing respective aspects of community resilience (Birkmann et al., 2014).

Although resilience indices are prominently featured in the literature, there are several documented issues. First, despite much overlap in their formulations and end goals, there isn't a clear way to compare what indices are comprised of or are actually measuring (Gillespie-Marthaler et al., 2019). Second and related, there is a lack of objective justification for the use of one indicator over another, which has reasonably led to a proliferation unproven indicators in the field (Fekete, 2019; Gillespie-Marthaler et al., 2019). Lastly, indices themselves are also questionable in their abilities to predict disaster outcomes that would intuitively align with notions of community resilience; Bakkensen et al. (2017) found that indices have mixed results when compared to expected outcomes of property damage, fatalities, and disaster declarations. In general, these issues reveal that the field is lacking construct validity, demarcations of the degree to which a model measures what it claims to be measuring. Until data-driven rigor is applied to investigating indices and their underlying components (i.e., indicators), researchers and practitioners will be unable to determine the extent to which elements of an index predict and/or causally explain disaster outcomes, thereby failing to holistically quantify community resilience.

Currently, objective comparisons are difficult to establish because of the sheer multitude of existing indicators, which often differ only trivially in what they're measuring, and the semantic ambiguity introduced throughout the indicator creation process that obfuscates underlying data. In other words, researchers do not have a clear and comprehensive set of independent variables from which to investigate construct validity. We believe this problem stems from the fact that the index creation process is predominantly informed by a top-down, theory-driven philosophy, as opposed to a bottom-up, data-driven mindset. The stipulation of capital systems, designation of indicators, and eventual selection of data, although reasonably constructed on an individual basis, are ultimately subjective in nature.

As such, to help pave the way for such empirical analyses and objective comparisons to take place, we deconstruct leading indices in the field into their underlying *variables*. We use the term variables to simply denote the measured set of features included an index, much like in a typical data analysis. For example, instead of stating there is an "infrastructure" capital system, comprised of a "medical care capacity" indicator, measured by "number of hospital beds per 1000" (Cutter et al., 2014), we simply say there is one variable, "hospital beds per capita". By starting at the data-level, we obviate the semantic ambiguity introduced by the indicators and capital systems.

Then, via factor analysis (FA), we establish an intuitive way to group the variables based on their statistical correlations measured at the US county level. Essentially, the analysis can be thought as an extensive data cleaning and dimension reduction problem to help clear up ambiguity in the indicator space. The resulting factor model comprises a data structure that researchers can effectively use as a baseline from which to conduct further empirical analyses, thereby helping to build construct validity in the field. Several indices have employed various correlation analyses similar to FA in their formations, but their focus has been on using these techniques as a means for determining how well indicators fit into capital system categories or aggregating indicators into an overall index (Cutter et al., 2014; Peacock et al., 2010; Sherrieb et al., 2010; Tapia et al., 2017). Instead, we borrow lessons learned from the psychometrics literature and suggest that the factor analysis itself should serve as an empirical baseline from which to measure community resilience.

It is important to note that with this analysis we are not dismissing the importance of theory-driven exercises in quantifying community resilience. On the contrary, we place great value in theory-driven exercises of how various capital systems and indicators can plausibly inform disaster resilience, which is exactly why we base this analysis on a comprehensive set of data derived from leading indices in the field. We are simply reorganizing the wisdom and expert knowledge embedded in the current literature in a manner that is more conducive to empirical validation. Like with many scientific endeavors, both theory- and data-driven philosophies are necessary for progress; currently, the community resilience literature is lacking the latter.

### 2.2 Background

At the 2005 World Conference on Disaster Resilience, it became clear that research on community resilience was lacking compared to society's ability to understand and predict natural hazards (UN/ISDR, 2005). Thus, efforts were made to better define and measure community resilience, and there have since been a proliferation of frameworks that purport to do so. Together, Cutter (2016); Sharifi (2016) have identified over 40 such approaches in the literature. In general, these frameworks entail either a localized case study application or a generalized geographic index (Cutter, 2016). Local case studies entail site visits and interviews for data collection and lend themselves to community engagement. Indices tends to rely on statistical analysis of widely accessible demographic and infrastructure data and cater more to policy development (Flanagan et al., 2011). Some studies incorporate a blend of both designs, such as Frankenberger et al. (2013) employing multi-level modeling techniques to merge household survey data within an index composition.

Another emerging phenomenon is the use of social media data for disaster preparedness. Social media allows for localized spatiotemporal data to be collected on a wide scale (Keim and Noji, 2011). Its potential for mitigation response has been well documented, and due to its ability to reflect direct human experience, has become another source of information to quantify community resilience (Chan, 2013; Niles et al., 2019; Wang and Zhuang, 2017). Recently, Rachunok et al. (2019) used Twitter data to populate resilience categories outlined in existing frameworks to examine patterns in messaging during events (Cimellaro et al., 2016; Cutter et al., 2008). Although we do not deal with social media directly in this analysis, nor are we aware of a geographic index that utilizes such data at this point in time, social media will clearly play a prominent role in the future of quantifying community resilience and offers another reason to lay a solid empirical foundation from which new sources of data can be assembled.

To the best of our knowledge, regardless of the approach, community resilience frameworks begin with a theory-driven stipulation of capital systems deemed pertinent to the analysis (Norris et al., 2008; Rowcliffe et al., 1999; Twigg, 2007). Then, various indicators, as defined in the previous section, are selected from the literature and/or crafted to fit into these categories. Next, the indicators are populated with corresponding data, either through on-site data acquisition in the case of localized frameworks or via demographic or other geographic information system (GIS) data in the case of indices. For local frameworks, some type of checklist and/or scorecard criteria are typically constructed. For indices, the data are generally standardized, subjected to dimension reduction techniques, and then aggregated into an overall index score (Cutter et al., 2003, 2014; Cutter, 2016; HVRI, 2016; Peacock et al., 2010; Sherrieb et al., 2010).

As mentioned, there are several issues with this process, and while the focus of this paper is on geographic indices, these problems also extend to localized approaches. There are over 40 community resilience frameworks and over 550 indicators identified in the literature (Cutter, 2016; Sharifi, 2016; Gillespie-Marthaler et al., 2019). However, there is no clear way to compare approaches, both in terms of what they are trying to measure and in terms of how well their indicators predict disaster outcomes. We surmise that these ambiguities are a shortcoming of the theory-driven philosophy that has predominated this space.

First, there is no consensus on how to define capital systems in the literature. For example, Renschler et al. (2010) proposed seven categories of capital systems for measuring resilience: people, environment, government, infrastructure, lifestyle, economic, and social, while Sherrieb et al. (2010) suggested four systems: economic, social, communication, and competence. There are a multitude of variations, differing in numbers and names of capital systems (Cohen et al., 2013; Cutter, 2016; Ebisudani and Tokai, 2017; Joerin and Shaw, 2011; Norris et al., 2008; Orencio and Fujii, 2013; Perfrement and Lloyd, 2015; Pendall et al., 2010; Shaw et al., 2010; Tapia et al., 2017; THRIVE, 2004; Yoon et al., 2016).

Furthermore, between indices, capital systems can be nominally the same but contain different indicators, or vice-versa, the same indicator can be included in different capital systems. For example, Peacock et al. (2010) include indicators of income, employment, property value, business, and health insurance in their economic capital system. Cutter et al. (2014) also designate an economic capital system but construct it with indicators of home ownership, employment rate, race and ethnicity income equality, non-dependence on primary and tourism sectors, gender income equality, business size, large retail store distribution, and federal employment. Conversely, Peacock et al. (2010) includes an indicator pertaining to households with access to vehicles within a physical capital system, while Cutter et al. (2014) include the same indicator as part of a social capital system.

This obscurity is exacerbated by the fact that there is no established method for populating indicators with data. Across indices, a similarly named indicator can be measured with different data, or vice-versa, indicators can differ in name but measure the same thing. For example, Peacock et al. (2010) measure an indicator for populations without health insurance based on all age groups, but Cutter et al. (2014) restrict this same nominal indicator to include only age groups under 65 years old. Conversely, the linguistic connection indicator named by Foster (2012) and language competency indicator named by Cutter et al. (2014) both use the same data. There are also numerous occurrences of similarly or differently named indicators that utilize only trivially different data. For example, Cutter et al. (2014); Foster (2012) have an indicator for the number of disabled people per total population, and Flanagan et al. (2011) have a similar one measured as the number of disabled people over five years old per total population.

While frameworks may be reasonably justified on an individual and theoretical basis, collectively, the lack of objective justification in defining capital systems, forming indicators, and selecting data makes comparing approaches difficult. Without communicating precisely what these indices are actually measuring and relating them to disaster outcomes, it is difficult to justify why one set of indicators is more appropriate than another for the task of quantifying community resilience. A natural consequence of this ambiguity and subjectivity is a proliferation of redundant, inconsistent, and unproven indicators that we find in the current literature.

At the overall index level, there have only been a few attempts at measuring performance against disaster outcomes. Most notably, Bakkensen et al. (2017) used several regression techniques to validate major indices against property damage, fatalities, and number of disaster declarations in the Southern US and found that indices have mixed results when compared to reasonable expectations of what should constitute as resilient communities. Peacock et al. (2010) regressed property damage and fatalities against their resilience index across counties along the Gulf Coast and found results aligned with expected outcomes, which helped give credence to their framework. Burton (2015) evaluated their resilience index based on a qualitative assessment of photographs taken during post-Hurricane Katrina recovery efforts, which was insufficient for precise insights but at least gave some directional assessment of their framework. These efforts are steps in the right direction, but until similar exercises are conducted for the comprising elements of the indices, construct validity in the field will ultimately remain absent; researchers will not be able to explain why one model is better than another or which features of the model are most important for predicting outcomes.

Our study attempts to use the results of a factor analysis conducted on a comprehensive set of variables derived from major resilience indices in the field to form an empirical foundation that will aid researchers in attempts to compare and validate indices and their elements. We argue that using a data-driven set of factors and variables offers researchers more transparency than the capital system and indicator designations, which in turn will aid efforts in conducting and communicating empirical validations. Additionally, Kotzee and Reyers (2016) have noted that community decision-makers found value in discussing the concept of resilience with respect to empirically determined components of their flood index, so it is reasonable that our approach will also make engagement with stakeholders more straightforward.

### 2.3 Methodology

This paper proposes a data-driven solution for comparing index-based frameworks attempting to quantify community resilience to natural hazards, thus providing a foundation from which researchers can quantify the extent to which measurable features in a community can predict disaster outcomes. In particular, we conduct an exploratory factor analysis on a GIS repository of US county level data based on a comprehensive set of variables from established resilience indices in the field. The resulting model forms a clear schema for comparing the constituent elements of these indices. A key benefit of the factor model is that it has the flexibility to incorporate new data and variables as needed while still providing a stable baseline for comparison and validation efforts.

#### 2.3.1 Approach

The current state of the literature on quantifying community resilience resembles that of psychometrics during the mid-20$^{th}$ century when psychologists were attempting to quantify personality. Similar to the concept of resilience, personality is multifaceted and varies extensively between applications (i.e., individuals), and numerous theory-driven models were developed in an attempt to measure it (John and Srivastava, 1999). Despite much overlap in frameworks, they were largely incompatible, and the few attempts at reconciliation resulted in more contention than clarification (McCrae and John, 1987).

Eventually, a data-driven approach using FA applied to surveys based on comprehensive sets of personality trait adjectives led to the development of a consensus model (Goldberg, 1992; John and Srivastava, 1999; McCrae and John, 1987). Linguistic relativity, the presupposition to this approach, hypothesized that because human interactions are intrinsic to societies across space and time, it is reasonable to assume that current language already comprises the words necessary to describe all facets of personality (McCrae and John, 1987). Thus, the dictionary provided an exhaustive yet finite list of features, from which inventories of trait adjectives were established and analyzed for common factors across subjects (Allport and Odbert, 1936; John and Srivastava, 1999). Here, we borrow insights from the psychometrics literature to drive an empirically-based schema for reconciling resilience indices. Analogous to the inventory of adjectives, we establish a representative set of variables for community resilience on which FA is conducted.

Using reviews performed by Cutter (2016); Sharifi (2016) as an initial reference point, we identified frameworks that met the following criteria: (1) they measure community resilience (and/or vulnerability) with respect to natural hazards in general, as opposed to type-specific disasters, (2) they are index-based approaches, allowing for intuitive statistical comparison across geographies, and (3) data used to construct them are publicly attainable and applicable at the US county level. These criteria were chosen because they encompass attempts to measure the general concepts of community resilience, as opposed to case specific applications, and help ensure the availability of quality data. Both requirements align with the stated goals of this paper, and for comparable reasons, Bakkensen et al. (2017) used a similar set of conditions when selecting indices to correlate against disaster outcomes. Additionally, Cutter et al. (2014) advocated the county level scale is most appropriate for spatial analysis because it is the most granular level at which demographic and physical data is consistently available, county governments are heavily involved in disaster mitigation activities, and county boundaries are fairly stable over time. Ultimately, our selection criteria led to the following six indices:

1. Baseline Resilience Index for Communities (BRIC) (Cutter et al., 2010)
2. Community Disaster Resilience Index (CDRI) (Peacock et al., 2010)
3. Community Resilience Index (CRI) (Sherrieb et al., 2010)
4. Resilience Capacity Index (RCI) (Foster, 2012)
5. Social Vulnerability Index (SoVI) (Cutter et al., 2003)
6. Social Vulnerability Index (SVI) (Flanagan et al., 2011)

BRIC is one of the most widely cited indices that attempt to quantify community resilience (Bakkensen et al., 2017; Cutter et al., 2010). Its stated goal was to establish a baseline set of criteria from which comparisons of resilience at the US county level can be made across time (Cutter et al., 2010). The 2014 version of BRIC, the one we utilized for this analysis, includes 49 variables (Cutter et al., 2014). Similar to the other five indices, BRIC's variables are informed by the general process discussed in the Background section: identify a wish-list of indicators to include based on theories of capital systems and their relationship to resilience, determine which of these indicators can be measured with corresponding data, and employ correlation analyses to winnow down the set of measured features (variables) (Cutter et al., 2014; Peacock et al., 2010; Sherrieb et al., 2010).

CDRI is another index attempting to measure resilience at the county level and includes 89 variables. However, in its original application, it only incorporated US counties along the Gulf Coast (Peacock et al., 2010). It should be noted that CDRI is one of the few frameworks to be validated against disaster outcomes upon its inception (Bakkensen et al., 2017; Peacock et al., 2010).

CRI attempts to "measure the sets of adaptive capacities for Economic Development and Social Capital in the Norris et al. (2008) community resilience model with publicly accessible population indicators" at the US county level (Sherrieb et al., 2010, p. 227). CRI includes 17 variables.

RCI is an index that measures resilience from economic and population stressors in addition to natural hazards (Bakkensen et al., 2017; Foster, 2012). It features 10 variables measured at the US county level.

SoVI is the prototypical framework for quantifying community resilience and/or vulnerability via an index (Cutter et al., 2003). SoVI attempts to determine the social capital of US counties, with the aim of revealing areas that might be lacking in related disaster mitigation capacities. The 2016 version, the one we utilized for this analysis, includes 29 variables (HVRI, 2016).

Lastly, SVI also attempts to quantify the social capital of communities (Flanagan et al., 2011). Unlike the other indices, SVI was developed using the US census tract as the spatial unit of analysis. However, all the underlying metrics are available and pertinent to the county-level, so we could reasonably include this index in our analysis. SVI consists of 15 variables.

Using the set of variables that span the six indices, we compiled a GIS repository of measurements at the US county level. This repository serves as the analogy to the set of trait adjectives established in the psychometrics literature. The comparable presupposition to linguistic relativity is not as strong here, as there are likely variables not included in these indices that can help quantify community resilience. Nonetheless, the dataset serves as an extensive set of currently measurable features at the county (or similar) level. Additionally, the dataset forms a clear baseline from which communities and decision-makers can be engaged to solicit new and potentially valuable variables for analysis. Similarly, others have mentioned that many of the indicators identified in the literature cannot yet be populated with data at this time (Cutter et al., 2003, 2010; Peacock et al., 2010; Sherrieb et al., 2010). This predicament may change in the future, and a key benefit of the FA approach, discussed in more detail later, is that it allows for a flexible incorporation of new variables.

### 2.3.2 Data

The six selected indices delineate data sources, fields, and transformations used to construct their corresponding variables, and we attempted to recreate these efforts with as much fidelity as possible with contemporary data. Because much of the underlying data stem from widely accessible and maintained sources, this task largely entailed us using the original variable definitions simply updated with more recent data. For example, SoVI relies exclusively on US Census Data, namely the American Community Survey and County Business Patterns tables, so we were able to completely recreate this index with current data (HVRI, 2016). Across indices, we were able to recreate the vast majority of variables. When we were not able to exactly reproduce a variable's definition, generally due to original data sources no longer being available, we attempted to find a suitable replacement metric. For example, CDRI includes a variable for the number of ambulances per population. The County Business Patterns no longer contains this information at the county level, so we used the number of ambulance service employees per population as a proxy. If finding a reasonable proxy was not feasible, we simply omitted the variable from the analysis. Overall, we were able to reproduce 180 of the original 204 (88%) variables specified across the six frameworks with recent data, including proxied variables. The missing 24 (12%) stemmed almost entirely from data sources no longer being available, with a small portion attributable to data that now costs money to obtain. Table 1 gives an overview of the variable coverage, and the Supplementary Materials contain a data dictionary that documents all details necessary to recreate this dataset, including notes on proxied variables when applicable.

Table 2.1: Index Coverage

| | Original | Reproduced |
|---|---|---|
| BRIC | 49 | 41 |
| CDRI | 84 | 72 |
| CRI | 17 | 14 |
| RCI | 10 | 9 |
| SoVI | 29 | 29 |
| SVI | 15 | 15 |

The 180 variables completely span 3,067 of the 3,142 (98%) US counties as defined by 2017 Census borders. Overall, the spatial coverage is expansive, and the vast majority of missing entries pertain to Alaska where recent changes in county borders have made them incompatible with some of the older data sources. With FA, it is necessary to account for these missing county entries. For the sake of transparency, we omitted the 75 (2%) counties that did not have complete coverage for the 180 variables. It should be noted that we also repeated the analysis by mean-imputing the missing entries, and the difference in results was negligible.

Additionally, we identified 83 variables that were redundant between indices. We define a variable to be redundant if it is constructed with identical or near-equivalent data. The vast majority of these redundant variables were exact duplicates, and the remaining were the near-equivalent cases, which we define to be variables that include trivially different data. For example, BRIC, SoVI, and SVI specify a variable that corresponds to age-dependent populations, but BRIC and SVI define this variable with an age range of over 65 years old while SoVI also includes populations under five years old. Ultimately, the redundant variables would provide no additional insights to the model and only bias the apportionment of variance contributed to the resulting factors. As such, we elected to remove exact duplicates and reduced the sets of near-equivalent cases into single features, the latter done by averaging derivations of the near-equivalent variables while maintaining compatible units. Ultimately, this data cleaning process yielded 130 unique variables on which we conducted FA. The Supplementary Materials include details and mappings of the redundant variables as well as an *R* programming script with all the data processing steps.

Lastly, prior to FA, we standardized the 130 variables using Z-score transformations. This standardization allows the FA model to be based on correlations between variables, not covariances which are unit-dependent. The Z-score transformation is common among multivariate analysis, including the creation of several resilience indices (Cutter et al., 2003, 2014; HVRI, 2016; Peacock et al., 2010; Sherrieb et al., 2010).

### 2.3.3 Factor Analysis

FA is a statistical dimension reduction technique that identifies latent factors accounting for variance common to observed variables (Revelle, 2018; Spearman, 1904; Thurstone, 1947). Here, we apply FA to the dataset of 130 standardized variables spanning 3,067 counties to find patterns of correlation among measures of community resilience at the US county level. The resulting model provides an effective means for comparing the six indices and lays a solid foundation for future validation attempts.

FA is often compared to a similar dimension reduction technique, principal components analysis (PCA). Both attempt to parsimoniously reconstruct the covariance or correlation matrix of the respective dataset by finding patterns that emphasize similarities between variables. However, PCA assumes that components (factors) are linear combinations of observed variables and finds an orthogonal basis onto which the variance of the projections of these variables is maximized across all the data (Revelle, 2018). The solution is deterministic, found via eigendecomposition or singular-value decomposition of the covariance or correlation matrix. FA, on the other hand, presumes that factors (components) are latent constructs whose linear combinations give rise to common aspects of observed variables (Revelle, 2018; Thurstone, 1947). Mathematically, this notion equates to FA solving for error terms on the diagonal of the covariance or correlation matrix that estimate variables' unique contributions to the variance (Revelle, 2018). Consequently, the total variance of the dataset becomes partitioned into unique and common portions, the latter of which is accounted for by the factors. FA is a generative model and cannot be analytically solved like PCA, instead requiring best-fitting algorithms to estimate a solution (Revelle, 2018).

Accordingly, for FA, the mapping of latent factors onto variables is not sensitive to the incorporation of new variables (Revelle, 2018). This notion does not hold true for PCA. Thus, by using FA as opposed to PCA, we can find a data-driven model that produces a stable foundation for comparing the constituent elements of the indices while also being flexible in the incorporation of new variables in the future. Granted, if many new and disparate variables are investigated, the factor model should be revisited to ensure that the number and interpretation of factors remain consistent.

Lastly, it should be noted that we applied varimax rotation to the loadings matrix of the factor solution in order to simplify its structure and make its interpretation more intuitive (Kaiser, 1958; Revelle, 2018; Thurstone, 1947). Loadings consist of the correlations between variables and components (factors), scaled by the square-root of the variance attributable to respective components (factors). Varimax rotation maximizes the variance of the squared loadings of all variables while still preserving orthogonality in latent space (Kaiser, 1958; Revelle, 2018). Varimax rotation does not affect the total variance explained by the model but simply partitions this amount differently between the factors (Revelle, 2018).

Several community resilience frameworks employ dimension reduction techniques. Cutter et al. (2003) used PCA combined with varimax rotation to address multicollinearity and remedy potential double-counting of variables in the creation of SoVI. Cutter et al. (2014) do similarly for BRIC. Pfefferbaum et al. (2015) employed confirmatory FA on survey results to help verify their model aligned with stakeholder intuitions of resilience. Similarly, Tapia et al. (2017) and Cutter et al. (2014) respectively used a Cronbach's alpha test and Principal Components Analysis (PCA) to compare empirical clusters of indicators with their predetermined capital systems; the former found empirical redundancies between capital systems, and the latter noted that resulting PCA clusters did not align with the capital system definitions. Clearly, the risk analysis literature has recognized the utility of these statistical techniques in helping to organize the vast and varied data available for quantifying community resilience. However, the focus has been on using these techniques to help create singular frameworks, as opposed to comprehensively reconciling frameworks as we advocate for here.

## 2.4 Results

### 2.4.1 Number of Factors

FA requires the modeler to determine the number of factors to be extracted from the variables in the dataset. There are several approaches for doing so. They can be driven by theory (see, Pfefferbaum et al., 2015), simply entail retaining factors until they become uninterpretable (Revelle, 2018), or be based on empirical methods such as a scree plot analysis, parallel analysis, and/or Kaiser Criterion. For our purposes, we wanted to reduce the dimensionality of the data as much as possible while still maintaining interpretable factors, in order to make validation efforts and engagement between and among researchers and practitioners more straightforward.

Figure 2.1 depicts a scree plot and parallel analysis. Here, eigenvalues, representing the total amount of variance explained by their respective factors, are charted in successive order. Characteristic of a scree plot test, factors beyond the "elbow", or leveling-off period of the curve, offer negligible contributions to the explained variance of the model (Cattell, 1966). In Figure 2.1, a reasonable case can be made for the retention of 5-8 factors using this criterion. When exploring the number of factors to retain in the model, we found those beyond the fifth one to be uninterpretable even with varimax rotation. Corresponding variable clusters and scores did not yield meaningful relationships, and as such, their inclusion would reasonably be counterproductive in helping to establish a consistent and practical baseline for comparing indices. Thus, we elected to retain five factors in the model.

As mentioned, it should be noted that other criteria can be used to determine the appropriate number of retained factors depending on the purpose of the analysis. A parallel analysis retains eigenvalues that are greater than those corresponding to a random sampling of the data, red-dashed line in Figure 2.1, in this case

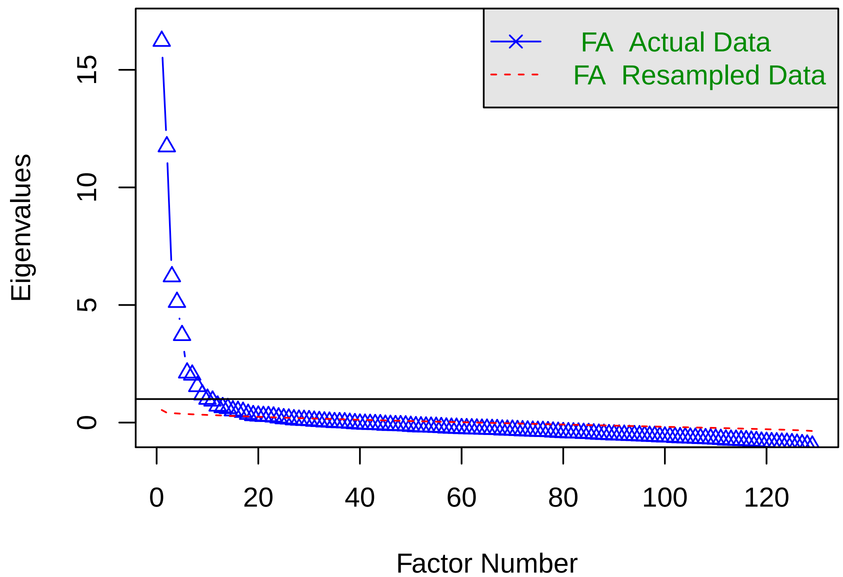



Figure 2.1: Scree Plot and Parallel Analysis

recommending 23 factors (Horn, 1965; Revelle, 2018). The parallel analysis criterion would be more appropriate if we were attempting to completely reconstruct the correlation matrix of the included variables. The Kaiser Criterion is a rule-of-thumb that retains all factors with an eigenvalue $\geq 1$, keeping factors that explain at least as much variance as a single variable does (Kaiser, 1960). Here, the Kaiser Criterion would suggest 10 factors should be retained, but as mentioned, factors beyond the fifth were uninterpretable. Ultimately, the question of how many factors to be retained is dependent on the scope of the analysis (Revelle, 2018). Although we found five factors most appropriate for our current task and dataset, if many new variables attempting to measure community resilience are investigated in the future, the model should be revisited and this number could change.

It should also be emphasized that an underlying assumption of FA is that factors are linear combinations of variables. If there are non-linear relationships between the variables attempting to measure aspects of community resilience, such as tipping points or interactions, FA would not account for them. However, the goal of our model was to simply reduce multicollinearity in the data, which in turn makes more complex investigations into variable interactions more straightforward in the future, as discussed in the Results section.

An overview of the current five-factor model with varimax rotation is shown in Table 2.2, and the full model output is available in the Supplementary Materials. The data and *R* scripts necessary to reproduce the

FA can also be found in the Supplementary Materials, should readers wish to explore different models.

Table 2.2: Factor Loadings

| Variable Description | Factor (Explained Variance) | | | | |
|---|---|---|---|---|---|
| | 1 (10.5%) | 2 (7.1%) | 3 (6.6%) | 4 (5.1%) | 5 (4.7%) |
| Median housing value | .854 | | | | |
| Median gross rent | .831 | | | | |
| %Households earning over $200,000 annually | .817 | | | | |
| %Population employed in creative class jobs | .815 | | | | |
| Per capita income | .791 | | | | |
| %Multi-unit housing | .741 | | | | |
| Median household income | .672 | -.640 | | | |
| %Asian population | .625 | | | | |
| %Population employed in legal services | .590 | | | | |
| %Landcover of pervious surfaces | -.566 | | | | |
| %Pop. employed as architects or engineers | .535 | | | | |
| #Architecture and engineering businesses | .524 | | | | |
| %Population living below poverty line | | .762 | | | |
| Food security rate | | -.736 | | | |
| %Children living in married couple families | | -.739 | | | |
| %Renters | | .675 | | | |
| %Owner-occupied housing units | | -.672 | | | |
| %Households with a vehicle | | -.653 | | | |
| Gender equality of median income | | .640 | | | |
| Inverse GINI coefficent | | -.598 | | | |
| %Single parents | | .596 | | | |
| %Black population | | .574 | | | |
| Inverse crime rate | | -.530 | | | |
| SoVI score | | .524 | | | |
| Businesses per capita | | | .750 | | |
| Retail stores per capita | | | .731 | | |
| Governments and special districts per capita | | | .621 | | |
| Urban influence score (rural higher) | | | .605 | | |
| Distance to metropolitan area | | | .603 | | |
| Religious organizations per capita | | | .531 | | |
| Road connectivity per capita | | | .527 | | |
| Hospitals per capita | | | .526 | | |
| Distance to nuclear power plant | | | .517 | | |
| Newspaper publishers per capita | | | .516 | | |
| Hotels per capita | | | .503 | | |
| %Population receiving social security | | | | .790 | |
| Median age | | | | .754 | |
| %Population $\geq$ 65 | | | | .753 | |
| %Unoccupied housing units | | | | .681 | |
| %Population under 65 and health insurance | | | | -.642 | |
| Housing density | | | -.503 | -.636 | |
| %Population under 18 | | | | -.526 | |
| %Population with a disability | | | | .509 | |
| %Population Spanish speaking | | | | | .749 |
| %Population ESL | | | | | .723 |
| Crowded housing | | | | | .698 |
| %Minority | | | | | .624 |
| %Population with health insurance | | | | | -.614 |

Table 2.2 (continued)

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| %Female householders | | | | | -.542 |
| %Voted in last presidential election | | | | | .507 |
| ... 80 other variables | | | | | |

* Factor loadings are not shown if they are < 0.5 (in absolute value)

As seen in both Figure 2.1 and Table 2, five factors cumulatively account for 34% of the variation across the 3,067 counties as measured by the 130 variables. 34% variance would certainly not be sufficient to constitute a factor model attempting to fully explain the correlations between all observed variables. However, we are simply using FA as a data reduction tool that is flexible in its ability to analyze new variables without having to redo the entire model. The five-factor model does indeed aid in this effort by greatly reducing the parameter space, with 50 variables notably loading (i.e., $|loading| \geq 0.5$) onto five factors. Ultimately, we have simplified the wisdom and knowledge embedded in leading indices attempting to quantify community resilience into an intuitive list of five factors, which account for 34% of variation between county-level measurements across the US, and 80 remaining variables. The results of this effort redress much of the redundancy and ambiguity in the current indicator selection process, helping to pave the way for validation efforts.

### 2.4.2 Factor Interpretations

Factor scores are displayed in Figure 2.2. These scores, in conjunction with the factor loadings shown in Table 2.2, are used to interpret the factors. It should be noted that we reversed the signs of Factors 2, 4, and 5 as displayed in Figure 2.2 and Table 2.2 compared to the full model output to more intuitively align with interpretations; with FA, the sign of the loadings is arbitrarily determined and only matters in context with respect to corresponding scores.

The first factor corresponds to wealth and accounts for 10.5% of the variation between counties. Variables that load highly onto this factor consist of predictors of wealth, namely high housing values, per capita incomes, and percentages of households earning more than $200,000 annually. Additionally, counties with the highest scores on the first factor, New York County (NY), San Francisco County (CA), and Fairfax County (VA), correspond to the wealthiest parts of the US. Via PCA, Cutter et al. (2003) also identified wealth as the primary component of inter-county variation among variables included in SoVI.

The second factor relates to poverty and explains 7.1% of the inter-county variance in the data. Scores largely feature historically impoverished counties in the Southeastern US. Variables with the notable loadings entail people living below the poverty line, lack of food security, unmarried parents, and Black populations.

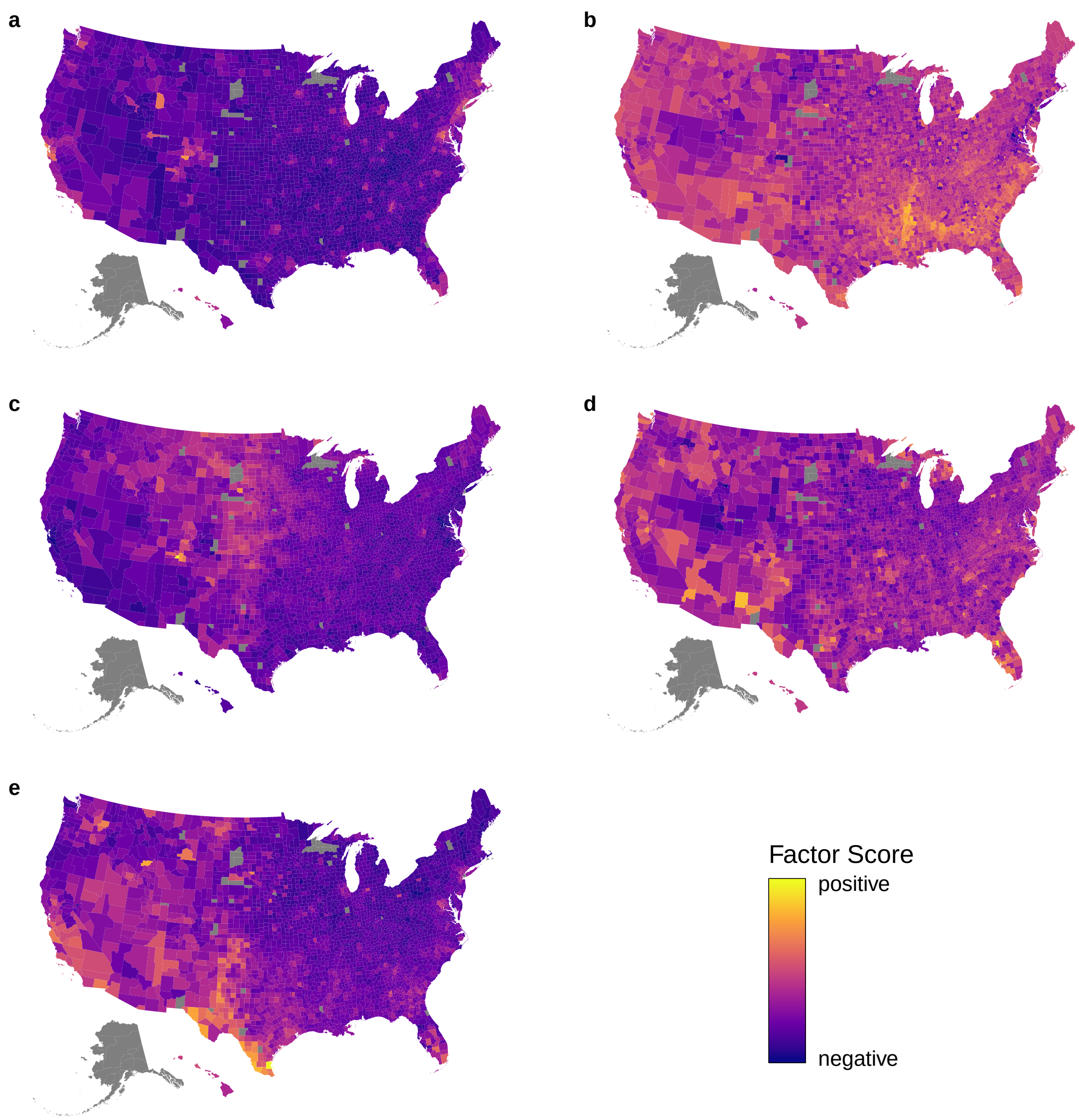



Figure 2.2: Community Resilience Factor Model: (a) Wealth, (b) Poverty, (c) Agencies Per Capita, (d) Elderly, (e) Non-English Speaking

Gender equality in median income is also correlated with poverty, as wage disparities on an absolute basis are minimized for lower overall incomes. It should be emphasized that poverty and wealth are related constructs but not simply opposites (i.e., the absence of poverty does not signify wealth), and this distinction is important to consider when discussing disaster preparedness. SoVI also featured a similar cluster of variables pertaining to poverty as its second largest component (Cutter et al., 2003).

The third factor accounts for 6.6% of the variance in the data and correlates with metrics that reveal the presence of a large number of business establishments, government institutions, religious organizations, and/or infrastructure developments relative to small populations. Scores highlight the agrarian Midwest US and unpopulated counties of Colorado rich in mining developments. In a sense, this factor appear to reflect a rural community's connectedness to the rest of world, both with respect to outside organizations (inter-connectivity) and internal developments (intra-connectivity). We term this factor as "agencies per capita", but its label is less important than attempting to understand its relationship to community resilience. Unique to the other factors, this one can reasonably be actively managed by community stakeholders, potentially providing a direct path to bolstering disaster mitigation. Additionally, most of the six indices do not include variables that load highly onto this factor, discussed in more detail in the following section, and may accordingly be lacking a valuable aspect of community resilience. The literature has already noted that measures of resilience differ between rural and urbanized areas, and a closer investigation of this factor via data analysis and community engagement would be a prudent exercise (Cox and Hamlen, 2015; Cutter et al., 2010; Rowcliffe et al., 1999; Tapia et al., 2017; Wilhelmi and Hayden, 2010).

The fourth factor relates to elderly populations. Variables loading strongly onto this factor pertain to older age groups, such as populations receiving social security benefits, populations with disabilities, and populations with higher ages. Additionally, featured areas include Florida, New Mexico, and Arizona, all among the most popular states for retirement in recent years (Folger, 2018). SoVI also found a component that pertained to elderly populations (Cutter et al., 2003; HVRI, 2016).

Lastly, the fifth factor involves non-English speaking populations. Variables with high loadings include predictors of English as a second language, Spanish speaking populations, and crowded populations. Scores highlight the Hispanic and Latino populations in Florida and states along the Mexican border. They also feature counties with high Native American populations in Montana and the Dakotas. Additionally, when we conducted the FA on the mean-imputed version of the dataset, this factor also included Alaska's indigenous population. SoVI also featured a similar component (HVRI, 2016).

It should be noted that median household income loads onto both the first and second factor, and housing density does similarly for the third and fourth factor. Results like these are expected and simply entail a variable being correlated with different clusters of variables. Varimax rotation attempts to minimize the occurrence of variables loading onto multiple factors, but it is rarely possible to avoid completely. Additionally, because loadings fundamentally measure statistical associations between variables, they do not necessarily have to be thematically related even though they often are. One example of a variable being correlated with others but not necessarily thematically related may entail the percentage of people who voted in the last presidential election loading onto the fifth factor; although, further investigation would be required to determine either way. Thematic explanations of factors aid in their interpretability and are based on the interpretable relationships of corresponding variables, but these relationships become less evident as factors account for progressively less variance and noise in the data begins to cloud underlying signals. This notion is precisely why we limited the model to five factors, as subsequent factors did not exhibit meaningful relationships between their corresponding variables.

### 2.5 Discussion

The main benefit of the factor model is that it transforms the expert knowledge embedded within a comprehensive set of resilience indicators into a simple, intuitive data structure that researchers can use to more easily conduct empirical inquiries. This notion applies to examining the coverage of existing indices as well as aiding in the investigation of a factor's or variable's ability to predict disaster outcomes.

Table 2.3 illustrates the former. For example, it is obvious that CRI lacks data pertaining to elderly populations, thereby revealing a potential gap in its data coverage. Additionally, CDRI is the only index that includes a variable covering ambulance services, and similarly, BRIC is the only approach that considers internet access with respect to disaster resilience. These types of data-driven comparisons are more precise and straightforward than discussions of capital systems and indicators. As such, researchers can more clearly communicate what features are already incorporated in the literature and curtail the spread of redundant and/or trivially differing indicators.

Additionally, because the factor analysis removes much of the collinearity in the data, the resulting model serves as natural set of independent features that researchers can use to empirically relate to disaster outcome data, thereby helping to build construct validity in the field. It is important to emphasize that just because the five factors capture the most variance between US counties for data that current literature deems is important for quantifying community resilience does not mean they are the best predictors of the such. In the spirit of Bakkensen et al. (2017); Peacock et al. (2010), we have begun the process of evaluating the five factors and 80 additional variables against disaster property damage and fatalities across the US using regressions and decision trees. These types of exercises will go a long way in helping to build construct validity in the field.

In a related sense, factor scores could also be investigated as random effects in mixed-effect models; for instance, using quintiles of Factor 4 scores as clusters could help reveal how main effects vary based

Table 2.3: Schematic for Comparing Indices

| | #Variables Loading onto Factors | | | | | |
|---|---|---|---|---|---|---|
| | BRIC | CDRI | CRI | RCI | SoVI | SVI |
| Fact1 (Wealth) | 1 | 6 | 2 | 1 | 5 | 2 |
| Fact2 (Poverty) | 5 | 4 | 3 | 2 | 6 | 3 |
| Fact3 (Agencies Per Capita) | 7 | 6 | 1 | | 2 | |
| Fact4 (Elderly) | 3 | 2 | | 1 | 5 | 2 |
| Fact5 (Non-English Speaking) | 2 | 3 | 1 | 1 | 4 | 3 |
| Var1 (Ambulance Service Employees per Capita) | | 1 | | | | |
| Var2 (%Population w/ Internet Access) | 1 | | | | | |
| Var3 (Building Construction Services per Capita) | | 1 | | | | |
| ... | | | | | | |
| Var80 (%Acres Wetland Buffers) | 1 | | | | | |

[*] We denote variables loading onto factors as those $\geq 0.5$ (in absolute value)
[†] Median housing value (SoVI and CDRI) loads onto both Factors 1 and 2
[‡] Housing density (SoVI and CDRI) loads onto both Factors 3 and 4

on the percentage of elderly population in a community. Similarly, the emergence of social media data can be spatiotemporally nested within the broader demographic data. For example, Zou et al. (2019) found that social-geographical disparities of Twitter usage exist throughout emergency management phases, and these types of insights and exercises could help predict outcomes when contextualized appropriately. The possibility of these types of empirical inquiries are immense and much needed in the current literature.

A natural extension to predictive inferences would be a researcher's desire to examine a factor's or variable's ability to mitigate disaster outcomes. These efforts involve examining causal relationships, which would require meticulous econometric approaches such as but not limited to difference-in-differences analyses or Bayesian structural time-series modeling. Having a clear and comprehensive set of empirical features makes these efforts much more straightforward. However, we want to emphasize that any attempts to establish causal relations between the five factors and disaster outcomes would require a more careful examination of the variables that are contained within these clusters, since they are only statistically associated and not necessarily causally linked. For example, even if Factor 4: Elderly Populations predicts increased disaster fatalities, perhaps only the disability rate variables causally explain these outcomes.

As such efforts to investigate predictive and causal relationships extend well beyond the means and findings of a single research group, the data-driven framework of the factor model allows for an orderly, organic backdrop from which to empirically validate notions of community resilience and communicate results. One researcher can determine how factors or variables predict a certain outcomes of interest, and his or her results can clearly be interpreted by the rest of the community without having to navigate through the semantic ambiguities introduced when defining capital systems and indicators. As described in Table 2.3, this approach provides a clear interpretation of what the scope of a study entails, helping to reveal potential information gaps and reduce trivial derivations between approaches. Additionally, because the factor model is flexible in its inclusion of new variables, researchers can also easily incorporate new data and variables without fundamentally altering the baseline structure. Granted, if many new variables are investigated, the model should be reexamined and updated accordingly as discussed in the methodology section. One could even blend the theory-driven mindset of designating capital systems and formulating indicators with this empirical baseline. For example, the theory-driven mindset could and should be used to inform additional variables, which in turn should be systematically validated against outcomes. Adopting this mindset would help curtail the spread of unproven and often trivially different indicators found in the literature.

Results of previous studies corroborate the viability of using the proposed factor model as a starting point for empirical investigations. SoVI found almost identical clusters of variables when addressing double-counting of measurements in its composite index (HVRI, 2016). Additionally, as mentioned, Kotzee and Reyers (2016) have already found that decision-makers placed value in discussing community resilience with respect to principal components. Additionally, Cutter et al. (2014) noted how PCA components derived in BRIC did not align with its theory-driven classification structure, calling into question the theory's a-priori authority and perhaps suggesting more flexibility in the creation process is needed. Overall, we find it reasonable that an intuitive and reduced set of factors and variables derived from an inclusive list of data from leading studies can provide a practical baseline for quantifying elements of community resilience and helping to establish construct validity in the field.

### 2.6 Consclusion

The concept of resilience explain why natural hazards of similar type and magnitude can have disparate impacts on varying communities. Numerous indices have been developed in attempt to measure community resilience. However, despite obvious overlaps in their formulations, there is currently a lack of continuity between approaches and a general lack of construct validity in the field.

To help address the former and pave the way for efforts to remedy the latter, we have developed a data-driven schema for comparing and validating index-based frameworks. Specifically, we conducted factor analysis (FA) on a comprehensive set of variables measuring aspects of community resilience at the US county level. The data entail 130 variables that span 3,067 of 3,142 US counties and are derived from six established indices that attempt to quantify community resilience and/or vulnerability: BRIC, CDRI, CRI, RCI, SoVI, and SVI. The FA suggests that 50 of the 130 variables effectively load onto five factors: wealth, poverty, agencies per capita, elderly populations, and non-English speaking populations.

The factor model reduces multicollinearity among the variables and provides an effective backdrop for comparing the constituent elements of indices, reasonably facilitating dialogue among and between researchers and practitioners and enabling validation efforts with disaster outcome data. Additionally, because the factors are insensitive to the inclusion of additional variables, the model serves as a stable yet flexible mechanism to determine whether newly proposed variables add value to the literature. As such, the risk analysis community can avoid a proliferation of redundant information as it investigates new measures for quantifying community resilience.

The five factors and other 80 variables need to be evaluated with respect to their ability to predict and possibly explain disaster outcomes. More specifically, although the five factors account for the most county-level variation with respect to the 130 variables, they may not be the best predictors of community resilience, much less offer policy makers a means of mitigating unfavorable disaster outcomes. However, until such empirical investigations are conducted in this field, efforts to quantify community resilience will ultimately remain incomplete.

### 2.7 Supplementary Materials

The data dictionary, dataset, and *R* scripts necessary to recreate the analysis are available on Open Science Framework: https://osf.io/vezns/?view_only=d8bdbceaa35e4acc9d14ada18f61fea4

# CHAPTER 3

## Applying Big Five personality traits and statistical learning to predict resource investment in community resilience

Community resilience plays a vital role in mitigating the risks associated with natural disasters. Public infrastructure that bolsters community resilience can often be conceptualized as a public good or common-pool resource. Here, we develop a common-pool resource game to investigate patterns in participants' contributions toward the development of public infrastructure that helps protect their community from various, hypothetical disaster scenarios. We train Bayesian additive regression tree (BART) models and multiple linear regressions (MLRs) on game results, given participants' Big Five personality traits and in-game circumstances as input variables, to predict deviations from strategies that would reasonably lead to Pareto efficient outcomes for the community. We find that in general, participants over-contribute relative to the Pareto efficient strategies, but the BART models are able to accurately predict these deviations and perform much better than their MLR counterparts. Additionally, we find several of the Big Five traits, Openness in particular, to have significant effects in predicting responses for both classes of models.

### 3.1 Introduction

Community resilience is a function of and often strengthened by public infrastructure and other related developments (e.g., sea walls, public hospitals, early storm warning systems). These resources are generally viewed as common-pool resources or public goods, depending if their utilization is considered to be rivalrous or non-rivalrous (Ayyub et al., 2016). For example, a community hospital benefits residents who are injured during a disaster but has limited capacity and as such is rivalrous in consumption (i.e., a common-pool resource). On the other hand, an early warning system for storms is non-rivalrous because its benefits are simultaneously shared by all residents. In either case, problems of collective coordination arise because although these developments bolster community resilience, their costs to individuals often exist as externalities (Ayyub et al., 2016). In other words, some community members may enjoy the benefits of these amenities without contributing individual resources to their development (i.e., free-ride).

The free-rider problem is not a new one. Hardin (1968) famously dubbed the term "Tragedy of the Commons" to describe the logical outcome of herdsmen overgrazing a pasture in pursuit of their own interests, but the nature of this problem dates back much earlier, likely to the dawn of life itself (Axelrod and Hamilton, 1981). Although the economically rational incentive for people is to free-ride in cooperative dilemmas, researchers find ample cases of successful, sustainable, and non-coercive management of communal resources in real life (Ostrom, 1999, 2000; Schindler, 2012). In fact, behavioral experiments in the form of economic games find that pro-social tendencies in cooperative dilemmas appear to be the norm, provided others do not defect from this pattern of behavior (Dong et al., 2016; Diekert, 2012; Schindler, 2012; Gunnthorsdottir et al., 2007; Capraro, 2013; Johnson, 2015).

There is a growing body of game-theoretic applications to disaster resilience (Seaberg et al., 2017). However, the majority of these studies assume decision-makers operate according to a rational or bounded-rational strategy and/or lack empirical justification for the use of employed decision models (Taylor et al., 2014; Bouzat and Kuperman, 2014). Here, we design a behavioral study in the form of a web-based common-pool resource game (CPR game) to empirically examine patterns in participants' contributions to the development of public infrastructure that helps protect their community from various disaster scenarios. We then analyze the results of the CPR game using statistical learning methods to predict participants' decisions given their in-game circumstances and personality traits.

### 3.1.1 Background

Economic games are a well-established approach to investigating decision-making patterns in social interactions. These games feature two or more individuals making decisions given specified conditions and information. Individuals' decisions not only affect their own outcomes but those of other participants. One class of games, called social or cooperative dilemmas, is characterized by a trade-off between short-term self-interests and long-term, collective interests (Zhao and Smillie, 2015). The Prisoner's Dilemma is a prominent example of such a game, where two participants decide whether or not to confess to a crime based on a given payoff matrix.

A public goods game (PGG) is essentially a multiplayer extension of the Prisoner's Dilemma. Players decide on a portion of their personal endowment to invest into a the public good, and the combined amount is multiplied by some return on investment factor, generally between 1 and the number of players, and then redistributed to all group members (Gunnthorsdottir et al., 2007). Collectively, the return on investment is maximized when all participants contribute, but individuals can improve their own returns if they withhold investments (i.e., free-ride). The dynamics of CPR games are similar to that of PGGs, but the communal resource is rivalrous in consumption and becomes depleted with non-cooperative behavior (e.g., The Tragedy of the Commons) (Hardin, 1968).

In general, the economically rational decision for individuals in the management of common resources is to pursue their own interests at the expense of others' (i.e., the Nash equilibrium is for individuals to free-ride) (Dong et al., 2016; Johnson, 2015; Capraro, 2013). However, researchers find that pro-social behaviors, defined as actions that deviate from the prescribed Nash equilibrium strategy to defect, are pervasive in both natural and laboratory settings (Ostrom, 2000; Dong et al., 2016; Schindler, 2012; Gunnthorsdottir et al., 2007; Kline et al., 2019; Diekert, 2012). This finding belies the assumption that people fundamentally act rationally in pursuit of their own interests (Dong et al., 2016). Much theoretical and empirical work has been dedicated to understanding the dynamics of this behavior (Johnson, 2015; Kline et al., 2019).

Because cooperation is often observed to decline as a game progresses, early investigations presumed that self-interests still lay at the heart of individuals' decisions and that individuals were simply trying to maximize long-term payoffs by initially cooperating (Dong et al., 2016). However, even in games lasting only one turn, many participants exhibit pro-social tendencies (Capraro, 2013). Subsequent studies indicate that various socio-psychological factors determine the success or failure of the management of a collective commons (Schindler, 2012; Ostrom, 2000). Individuals' attitudes towards concepts of reciprocity, trust, fairness, risk aversion, conformity, and/or inequity aversion can all affect outcomes (Schindler, 2012; Johnson, 2015). Furthermore, the relationships between socio-psychological factors and cooperative decision-making is complex, nonlinear, and often interdependent (Schindler, 2012; Poteete et al., 2010). For example, most subjects in incentivized PGGs tend to increase (or decrease) their contributions if they find they contributed less (or more) than others in the previous round (Dong et al., 2016). As such, by measuring individuals' differences in disposition, researchers may able to analyze characteristics that are predictive of certain response patterns in social dilemmas (Hirsh and Peterson, 2009).

A prominent method for quantifying socio-psychological dispositions of individuals is the Big Five model of personality (McCrae and John, 1987). Derived from the lexical approach in psychology in conjunction with factor analysis, the Big Five model is the most widely used and accepted method for measuring dimensions of personality in the field of psychometrics and has been shown to predict various patterns of behavior across a wide range of economic games (John and Srivastava, 1999; Zhao and Smillie, 2015; Kline et al., 2019). The model specifies five different dimensions (a.k.a., traits, factors, or domains) of personality, which are *Openness*, *Conscientiousness*, *Extraversion*, *Agreeableness*, and *Neuroticism*.

*Openness* reflects a disposition toward learning new things, trying new experiences, a lack of prejudice, and a willingness to trust others (Kline et al., 2019). As such, many researchers presume that higher Openness scores are predictive of pro-social tendencies in social dilemmas. In fact, several studies find a strong, positive relationship between Openness and cooperation in social dilemmas (Kline et al., 2019; Zhao and Smillie, 2015; Lönnqvist et al., 2011; Hilbig et al., 2013; Volk et al., 2011). The opposite trend has been found in only one cohort of one study Volk et al. (2012).

*Conscientiousness* characterizes individuals who are detail-oriented, organized, efficient, and industrious (John and Srivastava, 1999). This trait tends to be associated with fiscal conservatism and risk aversion (Zhao and Smillie, 2015). One might reasonably expect individuals who score higher on the Conscientious dimension of the Big Five to be able to more effectively follow a given strategy for the management of the commons, but the nature of this strategy is most likely determined by other traits and factors (Zhao and Smillie, 2015). Empirical evidence suggests as much, with only one study finding a significant relationship between this trait and pro-social behavior; in this case, it also happened to be a negative relationship (Ben-Ner and Kramer, 2011).

*Extraversion* reflects strong interpersonal tendencies and an enthusiasm for engaging in social activities (Kline et al., 2019; John and Srivastava, 1999). However, it can also be linked to more assertive qualities (Hirsh and Peterson, 2009). Therefore, it is reasonable to surmise that the more enthusiastic/affiliating aspect of Extraversion can positively predict cooperation in dilemmas, but perhaps, the more domineering side could do the opposite. The literature features contradictory effects, likely moderated by other factors in the game (Ben-Ner and Kramer, 2011; Brocklebank et al., 2011; Zhao and Smillie, 2015; Kline et al., 2019). For example, when participants are able to punish free-riders, Schroeder et al. (2015) find that Extraversion predicts cooperative behavior in PGGs, but Koole et al. (2001) find the opposite effect when this game dynamic is not present.

*Agreeableness* is characterized by an orientation to the needs of others, of being empathetic, considerate, and “nice”, and is thought to be the predominant trait underlying cooperation, fair play, and benevolence (John and Srivastava, 1999; Kline et al., 2019). Not surprisingly, studies find it to be the trait most predictive of pro-social behaviors across a wide range of economic games (Kline et al., 2019; Zhao and Smillie, 2015; Kagel and McGee, 2014; Pothos et al., 2011; Hilbig et al., 2013; Koole et al., 2001; Volk et al., 2011, 2012).

Finally, *Neuroticism* indicates a tendency to experience negative emotions such as stress, anxiety, and depression, especially when confronted by perceived threats, and to exhibit volatile or hostile behaviors (John and Srivastava, 1999; Kline et al., 2019). It is unclear how this dimension might influence behaviors in the social dilemmas, and again, the literature contains mixed results (Kline et al., 2019). Lönnqvist et al. (2011) find a strong, negative relationship in Neuroticism predicting cooperation in an incentivized Prisoner’s Dilemma, (i.e., participants get monetary bonuses based on performance in the game), but Hirsh and Peterson (2009) find the opposite is true for the withdrawal aspect of Neuroticism in a non-incentivized game where fear of being ostracized might play a larger role.

There is an emerging body of literature that uses personality traits to predict decision-making in cooperative dilemmas (Zhao and Smillie, 2015). One major challenge has been in assimilating and comparing results across studies, as each study can be geared towards specific applications and/or include numerous moderating effects (Zhao and Smillie, 2015). Some of these factors may include but are not limited to participants being able to punish free-riders, participants having access to the history of player actions, the degree of in-game communication allowed, in-game performance being monetarily incentivized, the framing of the dilemma, and the observed behaviors of others (Anderson, 2010). Recently, Kline et al. (2019) conducted a multi-level meta-analysis of 15 different studies in an attempt to hierarchically model these various factors and their interactions with Big Five traits. They find that Agreeableness and Openness are strong, positive predictors of pro-social behavior across these cooperative dilemmas (Kline et al., 2019). Although, to the best of our knowledge, the Big Five traits have not been used to model decision-making patterns in game-theoretic applications pertaining to community resilience.

There is a growing body of research that incorporates aspects of game theory and formal decision theory in research on disaster resilience (Seaberg et al., 2017; Rubas et al., 2006). However, the focus of these studies has traditionally been on modeling outcomes of disaster scenarios while assuming decision-makers operate according to some preset paradigm, instead of analyzing what might be the most appropriate decision models for the given application (Reilly et al., 2015). Additionally, most studies simply assume decision-makers act according to a rational or bounded-rational strategy and/or lack empirical justification for the use of the incorporated decision models (Zhuang and Bier, 2007; Adida et al., 2011; Bouzat and Kuperman, 2014; Zheng and Cheng, 2011; Taylor et al., 2014). Kunreuther and Michel-Kerjan (2015) are among the few who deviate from this trend and gather experimental evidence from a web-based game to model the likelihood of U.S. adults acquiring property insurance to mitigate flood risks.

### 3.1.2 Contributions

Here, we follow a similar approach to Kunreuther and Michel-Kerjan (2015) and design a web-based CPR game to examine interrelations among Big Five traits and decision-makers' behavior around opportunities to build public infrastructure that bolsters community resilience to disasters. While much work has gone into examining the effects of Big Five factors in game-theoretic applications, to the best our knowledge, no study has done so in the context of community resilience. Additionally, previous studies that utilize Big Five factors to predict behavior in cooperative dilemmas implicitly assume a linear relationship between these traits and response variables, despite researchers acknowledging that relationships between socio-psychological factors and decision-making are complex and nonlinear (Schindler, 2012; Poteete et al., 2010). Furthermore, decision-models that are used in game-theoretic disaster resilience applications generally lack empirical justification for their use and/or make simplifying assumptions thereof. Our work aims to help address these gaps in the following ways:

- We develop empirically-based decision models that can be used to predict or generate responses in CPR settings pertaining to public infrastructure investments that bolster community resilience to disasters
- We examine how the Big Five traits predict decision-making patterns in the above context and relate

findings to those established in other cooperative dilemmas

- We use Bayesian additive regression tree (BART) models to explore nonlinear relationships between Big Five traits and responses and compare results to those from multiple linear regressions (MLRs)

The remainder of this chapter is organized as follows. Section 3.2 describes our CPR game and the statistical learning approaches we used to model the results of the game. Section 3.3 details the results of the statistical learning procedures and relates findings to other studies that have used the Big Five factors to analyze behavior in cooperative dilemmas. Section 3.4 discusses possible extensions to our work, and Section 3.5 contains our concluding remarks.

## 3.2 Methods

Our methodology consists of two main components. First, we design a web-based CPR game where each turn, participants decide how much of their available funds they want to contribute toward developing public infrastructure that helps protect their community from an impending disaster. Second, we train BART and MLR models on the results of the game to predict decision-making patterns of participants, given their Big Five personality traits and in-game circumstances.

### 3.2.1 Community resilience CPR game

#### 3.2.1.1 Study overview

We implement our study using *oTree* in conjunction with *Amazon Mechanical Turk (MTurk)*. oTree is an open-sourced, Python-based platform for designing and implementing web-based tasks, such as single- and multi-player economic games and dynamic questionnaires, and has been referenced in over 600 academic publications (Chen et al., 2016). Amazon MTurk is an online crowd-sourcing platform that connects workers to small web-based tasks called Human Information Tasks or "HITs" (Sheehan and Pittman, 2016). MTurk offers academic researchers efficient and cost-effective access to a large, diverse subject pool, and as such, has been prominently featured in the machine learning, economics, and social sciences literature (Lee et al., 2018; Paolacci, 2010; Sheehan and Pittman, 2016). Studies find that compared to more conventional, in-person samples, MTurk participants are more demographically diverse and representative of the US population as a whole in terms of age, gender, race, education, and geography (Sheehan and Pittman, 2016). One of the main benefits of oTree is that it allows researchers to seamlessly publish tasks on MTurk, monitor the progress of participants, and provide worker compensation through its admin interface.

Overall, 100 participants successfully completed our study. To help ensure quality responses as recommended in *Amazon's Mechanical Turk for Academic Research*, we limited participation to workers who had at least 500 previously completed HITs with an acceptance rate of at least 95% (i.e., percentage of previously submitted work that was rated to be of quality submission) (Sheehan and Pittman, 2016). To avoid duplicate entries, we restricted the study to Amazon worker IDs that had not already participated in our study. Additionally, to avoid potential cultural barriers, we limited access to workers currently residing in the United States. However, we will likely relax the latter constraint in future studies. Participants were compensated $5 for completing the entire study, which takes approximately 20 minutes to finish. It should be noted that we did not award monetary bonuses based on the in-game performance of participants; this is an experimental design choice that we plan to vary in future work.

Our study consists of five main components: a consent form, a demographic survey, a Big Five personality assessment, a game tutorial, and the CPR game itself. The consent form is self-explanatory. The demographic survey consists of a standard questionnaire regarding a participant's race, gender, education level, employment status, age, and country of residence. Both can be found in the Supplementary Materials.

The Big Five assessment consists of the BFI-2-S inventory, a 30-item response that retains much of the reliability and validity as the 60-item BFI-2 inventory (Soto and John, 2017b). The BFI-2 inventory is a proven measure of Big Five personality traits that reduces acquiescence bias and improves accuracy compared to the established BFI inventory (Goldberg, 1993; Soto and John, 2017a). We elected to use the condensed version, the BFI-2-S, to help reduce the overall completion time of the study.

The tutorial explains the dynamics of our CPR game to participants by walking them through an example turn. Brief questions are interspersed throughout the tutorial to help ensure participants develop an understanding of the game prior to playing it, so we can improve our chances of receiving quality responses (Crump et al., 2013; Sheehan and Pittman, 2016). The entirety of the tutorial is included in the Supplementary Materials, but parts of it are highlighted in the body of the paper to explain the game's dynamics, detailed below.

#### 3.2.1.2 CPR game overview

In essence, our community resilience CPR game is a variation of the typical PGG. Participants act as members of a community that are going to be impacted by a sequence of disasters. Each turn, they must decide how much of their individual resources, measured in points, they want to contribute toward the development of public infrastructure that helps protect their community from an impending disaster (i.e., a "shield"). The combined investments are multiplied by a factor of two to determine the total size of the community shield. The standard PGG multiplier is between one and the number of players, and our game features three members per community (Johnson, 2015); however, both the multiplier and size of the community are factors we plan to vary in future studies. Since the shield has a finite capacity shared by the community, we technically categorize our game as a CPR game.

Each turn, players are presented with a disaster scenario; an example is shown in Figure 3.1. The scenario consists of four main parts: (1) the value of the community shield at the start of the turn, measured in points (e.g., 9 points), (2) the severity of the impending disaster, randomly generated on a scale from 1 to 50 (e.g., 10), (3) a chart showing each community member's available resources, represented as points (e.g., "You"in blue have 50 points remaining), and (4) a chart showing each community member's chance of being harmed by the disaster this turn, randomly generated from 0% to 100% (e.g., "You" in green have a 60% chance of being harmed). Note, although not displayed in Figure 3.1, the exact, numerical percentages of (3) and (4) are specified to players during the game. Also, in this version of the game, all players start with the same amount of resources, 100 points, and there are only three members in a community. In future studies, we plan to systematically explore the effects of varying these game parameters as well as what information is displayed to participants, such as making community member contributions known to participants.

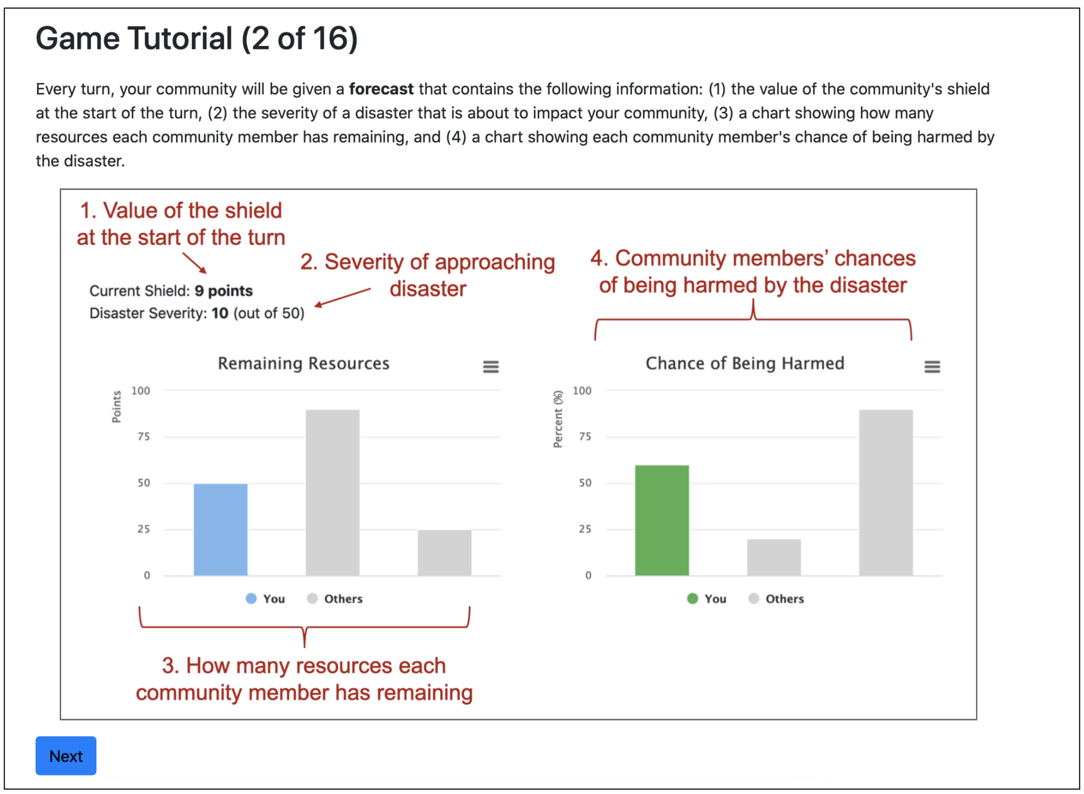



Figure 3.1: Game Tutorial (Page 2 of 16)

*Without* a shield, the amount of damage a player can expect to receive from a disaster is the severity multiplied by his or her respective chance of being harmed (e.g., without a shield, "You" would expect to receive 10 x 0.60 = 6 points of damage). The total expected damage to the community is just the sum of the expected damage to all its members (e.g., 6 + 2 + 9 points). The tutorial walks participants through these calculations with tangible counts, since people tend to more easily grasp the concept of expected value when it is presented in terms of frequency rather than probability (Gigerenzer, 1996).

Before the disaster causes any damage, participants can contribute some of their individual points to increase the value of the shield; participants cannot contribute more points than they have and cannot contribute fewer than zero points. In the example in Figure 3.1, the value of the shield at the start of the turn is 9 points. Assume that, when presented with the scenario in Figure 3.1, "You" contribute 2 points to the shield, "Player 2" contributes 0 points, and "Player 3" contributes 1 point. Therefore, the new value of the shield would be $9+2\times(2+0+1) = 15$ points. From previous calculations, the and then the next turn ensues in similar fashion. In our game, there are 5 total turns. Again, turn length is a factor we plan to vary in future work. If at any point a player runs out of resources, he or she will not be able to contribute any points to the shield for the remainder of the game. Currently, there is no option to punish others for not contributing, but we plan to implement this feature in future work.

Lastly, participants are given the option to contribute some of their resources before the game begins in order to increase the value of the shield before the start of the first turn. Participants are informed that the value of the shield at the beginning of the game is zero points. This game dynamic is meant to mirror real-life instances of people donating resources to relief efforts before they are needed, but this is another factor we plan to explore the effects of altering in future studies.

#### 3.2.1.3 Computer algorithms

In this version of the game, participants play with two different computer algorithms acting as the other members of the community. Participants are not given any information as to whether they are playing with other people or computers. The two algorithms are programmed to contribute amounts based on different strategies that would reasonably lead to a Pareto efficient outcome for the given turn.

The first strategy, which we call the "Pareto Efficient: Fair Share" strategy, instructs the computer to contribute an amount such that if all community members acted similarly, they would simultaneously reach a Pareto efficient for the group and individually efficient outcomes for themselves (i.e., they would contribute an amount that is proportional to how much they would benefit from the shield where the expected net loss to themselves and the community is both zero). For example, if there was *no starting shield* in Figure 3.1, the "Fair Share" strategy would entail "You" contributing 3 points, Player 2 contributing 1 point, and Player 3 contributing 4.5 points (i.e., each player's expected net damage divided by the 2x multiplier).

The second strategy, which we call the "Pareto Efficient: Equal Share" strategy, instructs the computer to contribute an amount such that if every community member contributed the same amount, irrespective of available resources and/or potential benefit from the shield, the group would arrive at a Pareto efficient outcome. In the same example *without* a starting shield, the "Equal Share" strategy would be for all community members to contribute $17 \div 3 \div 2$ points (i.e., total net damage to the community divided by the number of members divided by the shield multiplier).

One computer-controlled community member operates according to the first strategy, and the other, the second. The computers maintain the same strategy throughout the simulation, regardless of the decisions made by participants. However, we introduce some slight, random variation to the actual contributions made by the algorithms, so participants can't detect exactly how the computers are playing. On a related note, contributions are limited to whole numbers, so in general, the strategies would be close to but not exactly Pareto efficient solutions even without the random variation.

We employ these two algorithms for several reasons. Primarily, we wanted the default decision-making tendencies exhibited by other community members and observed by participants to be some form of pro-social behavior. The literature shows that people tend to cooperate in social dilemma games as long as they perceive other players to be doing similarly, and we wanted to explicitly control this dynamic in this version of our game (Ostrom, 1999; Johnson, 2015; Dong et al., 2016). Both the "Fair Share" and the "Equal Share" strategies are forms of cooperative behavior. Moreover, these strategies are intuitive to everyday notions of what constitutes pro-social behavior (e.g., when splitting the dinner check with friends, both "Fair Share" and "Equal Share" strategy are generally deemed to be socially acceptable). Additionally, we calibrated the game such that when the computers acted according to these strategies, the game would be balanced in terms of difficulty for the participants. For example, if the computers didn't contribute any resources to the shield, the disasters would regularly wipe out at least one community member by the time the game ends. Similarly, if the computers massively over-contributed to the shield, players would almost never have to worry about receiving any damage from the disaster. We plan to feature different algorithms in future studies. Similarly, we aim to run variations of the game where participants play with other people, instead of with computers, or even with "themselves" (i.e., have computers mirror participants or play according to a similar Big Five personality profile).

### 3.2.2 Statistical Learning

#### 3.2.2.1 Data

The results from the behavioral study comprise the data on which we train statistical models to predict decision-making patterns of participants, given participants' Big Five personality traits and in-game circumstances. Overall, 100 participants completed the study, and with a game lasting 5 turns, this amounts to a 500 total decisions. We reviewed the length of time participants spent on each part of the study to help determine if they took the study seriously or if responses should potentially be discarded. Similarly, for any outlying data points (i.e., contributions that were abnormally high), we reviewed how the corresponding participant played the rest of the game and/or if he or she left optional feedback at the end to gauge whether his or her decisions were genuine. Ultimately, we retained responses from all 100 participants. The Supplementary Details section contains the data from the game.

Additionally, we examined the hierarchical nature of the data (i.e., decisions nested within participants) when considering what statistical methods were appropriate to use. Specifically, we calculated the Intraclass Correlation Coefficient (ICC) to determine if multilevel modeling was warranted (Koo and Li, 2016). Overall, the ICC was very small ( 2.2%), informing us that variation explained at the subject-level was much smaller than that at turn-level. As such, a nested modeling approach was not necessary.

*Dependent variables* consist of two measures: deviations (i.e., differences in point values) between participants' contributions each turn and their theoretical contributions if they were operating according to the previously discussed "Fair Share" and "Equal Share" strategies. The range for both of these measures is approximately between -20 and 95 points. We elected to use deviations from these strategies as dependent variables because we found greater utility in discussing results in terms of over- or under-contributions relative to Pareto efficient strategies rather raw contribution amounts. Additionally, we repeated our analysis using raw contributions and contributions expressed as percentages of participants' resources, and these models resulted in slightly worse fits but fundamentally the same takeaways (see, Supplementary Materials).

*Independent variables* consist of fifteen inputs, of which six are modeled at the participant-level and nine are modeled at the turn-level:

- Participant-level: *Openness*, *Conscientiousness*, *Extraversion*, *Agreeableness*, *Neuroticism*, and pregame contribution

- Turn-level: player resources, other member #1's resources, other member #2's resources, shield value, expected damage (player), expected damage (group), potential savings (player), potential savings (group), and turn number

Participant-level variables entail BFI-2-S scores for each Big Five trait and any pregame contributions made to the community shield. We examined the potential role of having pregame contributions be a response or input variable, and ultimately decided on the latter. We found that the Big Five traits, in of themselves, were not sufficient to predict pregame contributions, and furthermore, studies show that initial contributions in social dilemmas can be good predictors of subsequent actions (Wunder et al., 2013; Nay and Vorobeychik, 2016).

Turn-level variables consist of several items, including each community member's available resources. Practically, these inputs directly relate to one's ability to contribute in the game, and studies have also shown that both an individual's resources and others' resources affect decision-making (e.g., some people are motivated by aversion to inequity) (Zhao and Smillie, 2015). Other inputs are the starting value of the shield, expected raw damage to the participant and community (i.e., the magnitude of the disaster), and the potential savings of the participant and community by investing in the shield (i.e., the net reduction in damage). Lastly, turn number is also included as an input variable because previous work has shown that cooperation tends to decline as the cooperative dilemma progresses (Dong et al., 2016).

It should be noted that we repeated our analysis using responses from the demographic survey as additional participant-level input variables. However, we found that these demographic variables were consistently among the least most important variables and added negligible explanatory power to the BART models (see, Supplementary Details). As such, we elected not to use these variables for simplicity's sake.

It should also be noted that several experimental design choices are implicit, game-level treatment factors in our models. These factors include but not limited to the following items: the size of the community, whether a participant is playing with humans or computers, what type of strategies or behaviors other community members exhibit, the value of the shield multiplier, the game length, the distribution of starting resources, the method of redistributing shield benefits, and the information presented and available to participants. We plan to systematically vary and explore the impacts of these treatments as game-level predictors in future studies.

#### 3.2.2.2 Bayesian additive regression tree (BART) models

We use BART models to predict player deviations from the two Pareto efficient strategies. Unlike parametric regression which imposes stringent assumptions on relationships between variables, BART instead non-parametrically learns these dynamics directly from the data (Chipman et al., 2010; Prüser, 2019; Hill et al., 2020). BART is a fully Bayesian implementation of a sum-of-trees regression model that has achieved critical acclaim in a wide variety of applications spanning multiple disciplines (Chipman et al., 2010; Kapelner and Bleich, 2016; Hill et al., 2020). Given an $N \times 1$ vector of response variables $Y$ and an $N \times p$ matrix of of predictor variables $X$, BART can formally be expressed by the following equation:

$$Y_n = \sum_{j=1}^{m} g(x_n; T_j, M_j) + \varepsilon_n, \quad \varepsilon_n \sim \mathcal{N}(0, \varepsilon^2) \tag{3.1}$$

Here, we have $m$ decision trees, each composed of a tree structure $T$ that consists of interior node decision rules and terminal nodes, and a set of values associated with each of the terminal nodes $M$ (i.e., leaf values) (Prüser, 2019; Hill et al., 2020). The interior node decision rules are binary partitions of the predictor space, $\{x_i <= c\}$ vs. $\{x_i > c\}$ (Chipman et al., 2010). Within a tree $T$, each observation is assigned to a terminal node based on the sequence of interior node decision rules from top to bottom of the tree (Chipman et al., 2010; Hill et al., 2020). In this manner, the tree structure captures complex, nonlinear interactions between variables (Prüser, 2019). However, decision trees in general can be prone to over-fitting, so BART places regularization priors on the size and fit of each tree-parameter pair $(T_j, M_j)$ so that a given tree contributes only a small portion to the ensemble's overall fit. In turn, the final prediction for each observation $Y_n$ is the sum of its corresponding leaf values across the tree ensemble.

BART priors consist of three main components: (1) the tree structure $T$, (2) the terminal node values $M$, given the tree structure, and (3) the error variance $\varepsilon^2$ (Kapelner and Bleich, 2016). BART models encourage small trees (i.e., low depth) that explore a range of input variables and splitting criteria. However, they also incentivize the terminal node values to be materially zero, in turn contributing little to the overall fit unless the tree-parameter pair $(T_j, M_j)$ results in a meaningful partition. As a consequence, BART features built-in variable selection, making it well suited for exploring a high-dimensional predictor space (Prüser, 2019). Chipman et al. (2010) discuss exact details on the specification of these priors. Moreover, because the tree ensemble includes a variety of tree structures, the final prediction inherently includes the main effects of the relevant predictors as well as any meaningful interactions between them, which are often unknown or too complex to specify parametrically (Chipman et al., 2010).

Although BART's default hyper-parameters for the priors are generally robust to different data sets and applications, we still perform 5-fold cross-validation to select the models with the lowest out-of-sample Root-mean-square error (RMSE) (Kapelner and Bleich, 2016). Given the observed data and specification for the aforementioned hyper-parameters and number of trees, BART then uses Monte Carlo Markov Chain (MCMC) techniques to perform iterative back-fitting of the tree ensemble to generate a posterior distribution of predictions (Chipman et al., 2010).

From this posterior distribution, it is straightforward to estimate the partial dependence function of input variables. Partial dependence functions, also known as partial dependency plots (PDPs) when in pictorial form, summarize the marginal effects that a specified set of input variables has on the response variable, while properly controlling for the effects of other inputs (Friedman, 2001). PDPs are analogous to parameter coefficients in linear regression, except they can be construed for any statistical model and make no a-priori assumptions regarding their relationship to the response variable. When working with complex, non-parametric models such as BART, PDPs provide researchers an intuitive means of examining the effects various input variables have on predictions. The formal expression for the partial dependence function is given in equation (3.2).

$$\hat{f}_{x_S}(x_S) = E_{x_C}\left[\hat{f}(x_S, x_C)\right] = \int \hat{f}(x_S, x_C)p(x_S)d(x_C) \tag{3.2}$$

Here, $x_S$ is the set of input variables of interest, and $x_C$ is the complement set. By integrating, or marginalizing, over the complement set of inputs, one obtains a function that depends only on the variables of interest while still accounting for their interactions among all input variables (Friedman, 2001). The partial dependency function is estimated via Monte Carlo approximation, as shown in equation (3.3).

$$\hat{f}_{x_S}(x_S) \approx \frac{1}{n}\sum_{i=1}^{n}\hat{f}(x_S, x_C^i) \tag{3.3}$$

In essence, the PDP is calculated by taking the average prediction across the various values for $x_C^i$ in the data while holding $x_S$ constant. Values for $x_S$ can also come from the data or be specified as a grid of points.

Lastly, we compare results from the BART models to those from MLR counterparts. The MLR models feature the same input and response variables as the BART models, and we do not specify any interaction effects. As mentioned, previous studies that have related Big Five factors to decisions in cooperative dilemmas employed some form of linear regression to do so. By comparing insights drawn from the BART models to those from the MLRs, we can assess if previous findings should be revisited with more advanced, non-parametric statistical approaches.

### 3.3 Results

Cronbach's alpha coefficients in Table 3.1 confirm that the BFI-2-S inventory serves as an acceptable measure for the Big Five traits in our study, with each coefficient measuring at least 0.70 (Cronbach, 1951; Shabrina et al., 2018). Given these BFI-2-S scores and the 10 other input variables, the BART models are able to accurately predict deviations from the "Pareto Efficient: Fair Share" and "Pareto Efficient: Equal Share" strategies, as seen in Figure 3.2.

Table 3.1: Cronbach's Alpha for BFI-2-S

| Openness | Conscientiousness | Extraversion | Agreeableness | Neuroticism |
|---|---|---|---|---|
| 0.75 | 0.77 | 0.72 | 0.70 | 0.81 |

The BART models perform much better than their MLR counterparts. Both classes of models tend to under-estimate the magnitude of large under- and over-contributions relative to each strategy, but this bias is much less pronounced in the BART models. As seen, the 95% posterior credible intervals for both BART

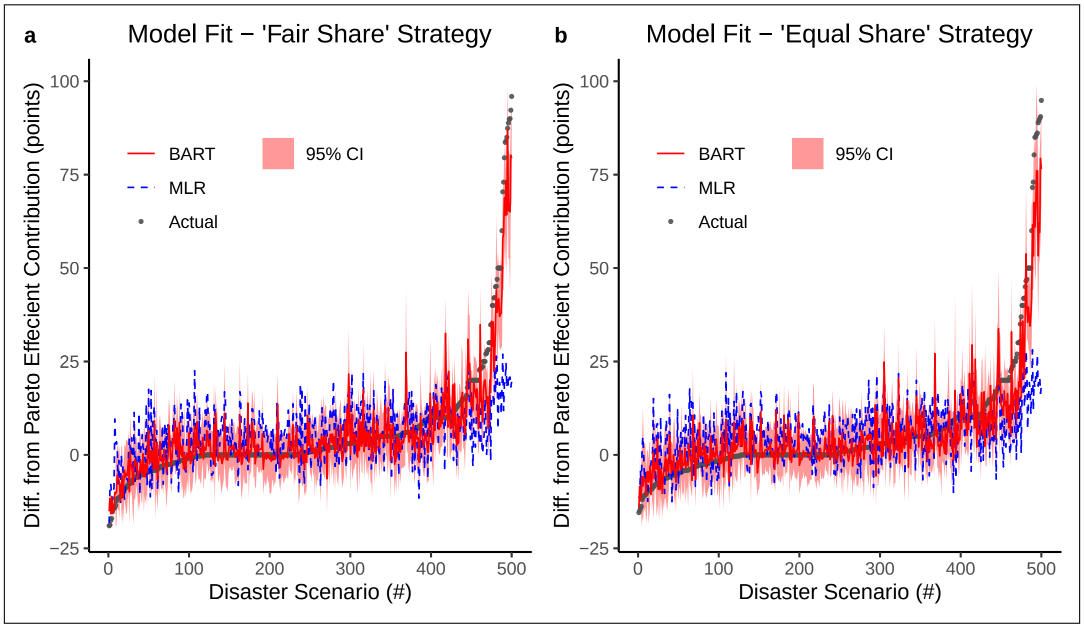



Figure 3.2: BART Model Predictions

models span the majority of responses. RMSEs for the mean BART predictions are approximately 5.94 and 6.00 for the “Fair Share” and “Equal Share” strategies respectively, compared to 14.24 and 14.25 for the corresponding MLRs. Also noteworthy is the fact that participant deviations from either strategy are significantly greater than differences in contributions between the two strategies, meaning that many participants’ decisions do not follow either of the Pareto efficient strategies.

Variable importance for BART model predictors is shown in Figure 3.3. Variable importance is the proportion of times a predictor is chosen as a decision rule out of all the decision rules in the post-burn-in posterior samples of the given BART ensemble (Kapelner and Bleich, 2016). This metric does not measure the effect a given predictor has on the response variable but is a measure of how often each predictor is used in each model’s formulation. For both BART models, a participant’s starting resources and scores for traits Openness, Extraversion, and Neuroticism and the game’s turn are featured in many of the decision rules. Interestingly, a participant’s potential savings from the shield is used more frequently in the “Fair Share” model compared to the “Equal Share” model, and vice-versa for the group’s potential savings (i.e., the model that attempts to learn a more individual-based strategy uses the more individual-based predictor for savings, and the opposite holds true for the more communal-based strategy). Note, we examined the correlation matrix of input variables to make sure multicollinearity was not a potential issue with variable selection in these algorithms (see Supplementary Materials).

Next, we perform a statistical test to determine which predictors in the BART models have significant

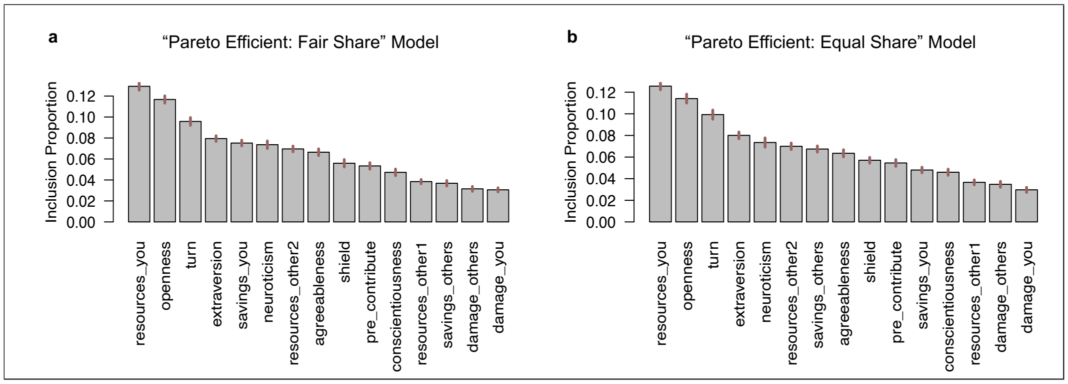



Figure 3.3: BART Variable Importance

effects on responses while controlling for the effects of other variables. We first permute the $x_j$th predictor of the dataset, which removes any relationship between $x_j$ and the response, and then compare samples of a pseudo-$R^2$ resulting from this model to that of the original BART model. Here, pseudo-$R^2$ is defined as $1 - \sum_{i=1}^{n_{train}} \varepsilon_i^2 / \sum_{i=1}^{n_{train}} (y_i - \bar{y})^2$ (Kapelner and Bleich, 2016). In turn, a quasi $p$-value is calculated as the average proportion of pseudo-$R^2$s across multiple permutations that are greater than that of the original model (Kapelner and Bleich, 2016). The null hypothesis is similar to that of a one-sided $p$-test for regression models, but the constraints for linearity are removed. Table 3.2 shows the results of these statistical tests and compares them to traditional, two-sided $t$-tests for the MLR counterparts. Note, Table 3.2 does not include intercept terms for the MLRs, but they can be found in the Supplementary Materials and are found to not significantly differ from zero.

Table 3.2: Predictor Significance

| | BART ($Pr > p$) | | MLR ($Pr > \|t\|$) | |
|---|---|---|---|---|
| | "Fair Share" | "Equal Share" | "Fair Share" | "Equal Share" |
| openness | $<$ **0.001** | $<$ **0.001** | $^\dagger <$ **0.001** | $^\dagger <$ **0.001** |
| conscientiousness | 0.225 | 0.297 | $^\dagger$**0.009** | $^\dagger$**0.009** |
| extraversion | **0.019** | 0.059 | $<$ **0.001** | $<$ **0.001** |
| agreeableness | 0.103 | 0.059 | $^\dagger$**0.009** | $^\dagger$**0.009** |
| neuroticism | 0.069 | 0.050 | 0.195 | 0.195 |
| resources_you | $<$ **0.001** | $<$ **0.001** | **0.001** | **0.001** |
| turn | 0.376 | 0.307 | $^\dagger$0.116 | $^\dagger$0.116 |
| savings_you | 0.189 | 0.356 | $^\dagger <$ **0.001** | $^\dagger$0.074 |
| resources_other2 | 0.050 | 0.050 | $^\dagger$0.131 | $^\dagger$0.131 |
| shield | 0.245 | 0.267 | 0.986 | 0.986 |
| pre_contribute | 0.236 | 0.228 | 0.268 | 0.268 |
| resources_other1 | 0.256 | 0.248 | 0.107 | 0.107 |
| savings_others | 0.584 | 0.327 | 0.768 | $^\dagger$0.145 |
| damage_you | 0.374 | 0.396 | **0.017** | **0.017** |
| damage_others | 0.404 | 0.347 | $^\dagger$0.754 | $^\dagger$0.754 |

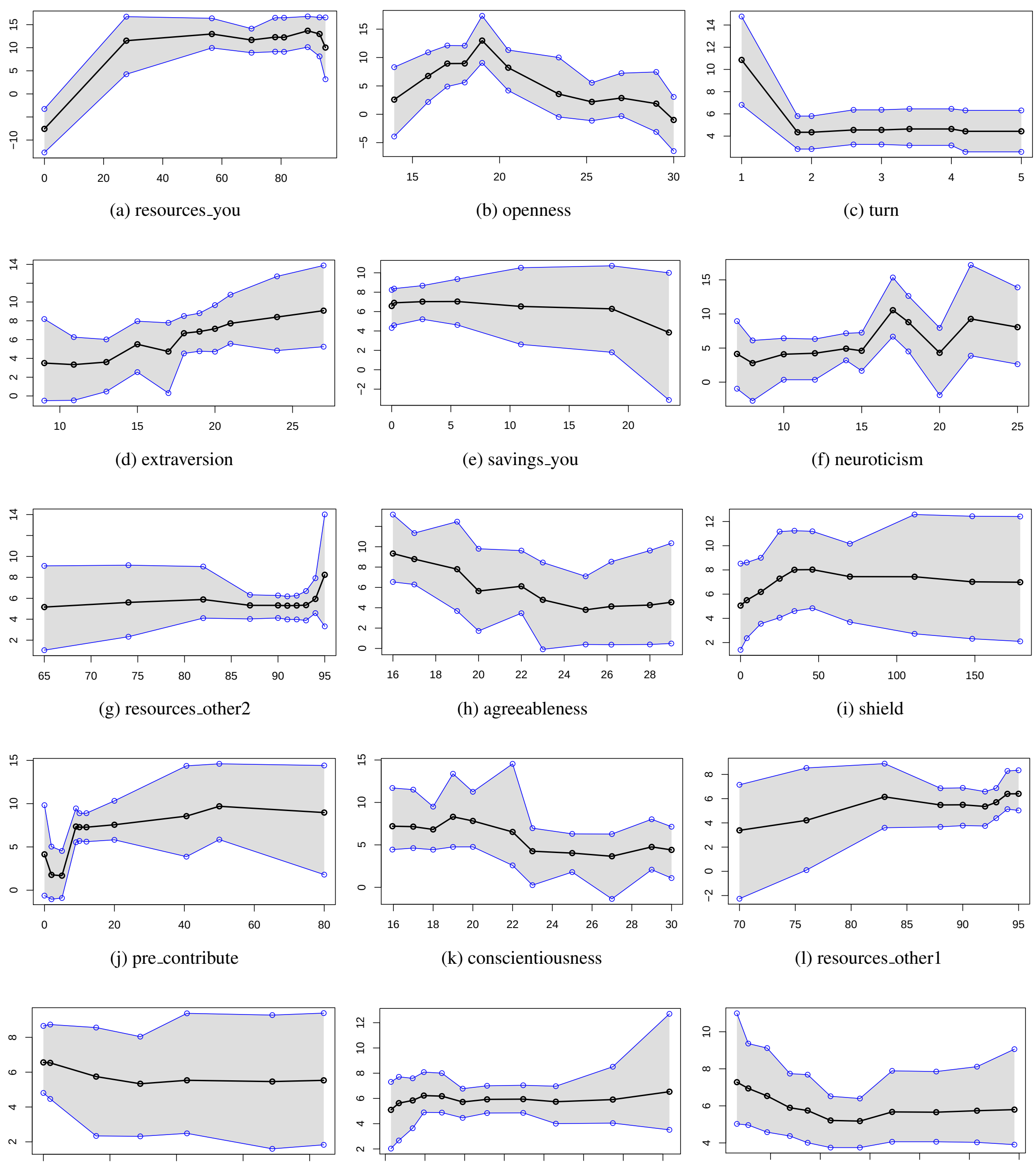
$^\dagger$ negative effect

The effects predictors have on responses are fairly similar, both within and between model classes. In general, the effects of the Big Five traits are prominently featured, and trait Openness has the lowest *p*-values (*t*-values) out of all the input variables. Additionally, a participants' available resources is also a significant predictor of responses in all models. However, one key distinction to make between model classes is that the BART models do not have negative or positive effect designations; this is due to the fact that unlike in MLR, predictors in the BART models are not assumed to be linearly related to response variables. Instead, their corresponding effects on responses are best viewed through PDPs, depicted below in Figure 3.4.

Figure 3.4 shows the posterior mean and 95% credible intervals of the partial dependency functions of all 15 input variables pertaining to the "Fair Share" BART model, displayed in order of this model's variable importance (see, Figure 3.3). PDPs for the "Equal Share" BART model can be found in the Supplementary Materials but are similar to those shown in Figure 3.4; additionally, we discuss any notable differences in the ensuing paragraphs. The PDPs are calculated at the deciles of the given variable of interest. The y-axes in the charts in Figure 3.4 measure deviations, in points, from the "Fair Share" strategy. Thus, values at $y = 0$ represent decisions that follow this strategy exactly.

Figure 3.4a shows that the lowest decile of participants' available resources predicts an under-contribution relative to the "Fair Share" strategy. However, for the remaining deciles, there is a consistent over-contribution. When players have little to no resources available, they are likely unable to contribute what is necessary to reach the given Pareto efficient decision, so the marginal effect is below $y = 0$. However, in general, we see the majority of responses are well above the Pareto efficient outcome, which may have ties to the concept of *risk-aversion* in decision theory (discussed later).

In Figure 3.4b, we see trait Openness predicts little to no deviation from the "Fair Share" strategy at low and high scores but positive deviations at middle deciles. Openness reflects an orientation toward trusting others and has been positively associated with pro-social behavior in social dilemmas (Kline et al., 2019; Lönnqvist et al., 2011; Hilbig et al., 2013). As such, the first half of the Figure 3.4b is intuitive in that we could reasonably expect to see a greater propensity for over-contributions relative to the "Fair Share" strategy in participants with higher Openness scores. However, trait Openness also pertains to intellect, associated with general intelligence and/or an ability to solve complex problems. Therefore, it is also reasonable that the highest Openness scores are predictive of participants who play the game closer to a Pareto efficient strategy (i.e., they're still engaging in pro-social behavior, just in an efficient manner). The Big Five Aspect Scale (BFAS) inventory, created by DeYoung et al. (2007), measures these two distinct but correlated aspects of Openness and could be used here to help confirm this notion.



(a) resources_you (b) openness (c) turn

(d) extraversion (e) savings_you (f) neuroticism

(g) resources_other2 (h) agreeableness (i) shield

(j) pre_contribute (k) conscientiousness (l) resources_other1

(m) savings_others (n) damage_you (o) damage_others

Figure 3.4: BART “Fair Share” PDPs (point deviations vs. deciles)

Figure 3.4c shows that the game's turn predicts large over-contributions relative to the "Fair Share" strategy for the first turn and less pronounced but still consistent over-contributions for the remainder of the game. The literature has shown that pro-social behavior tends to decline as cooperative dilemmas progress, so the PDP could partly be reflecting this dynamic (Dong et al., 2016). However, much of the rationale for this decline in social behavior is the result of defectors causing others to become uncooperative. In our game, some form of cooperative behavior is always exhibited by the computers, so unless participants feel disenfranchised when the algorithms contribute zero points to a community shield that is already able to withstand the impending disaster, there is likely some other motivation behind this behavior. This notion is corroborated by the fact that the PDP is still positive throughout the duration of the game (i.e., participants are still over-contributing in later turns, just to a lesser degree). More likely, participants either do not fully comprehend game at first and contribute much more than what is necessary, or they are knowingly sacrificing their own resources to ensure the rest of the community is protected. The pregame tutorial helps mitigate the risk that participants do not understand the game, albeit to an imperfect degree, and optional, post-game feedback indicates that most of the participants understood the game. However, in future studies, we plan to include another post-game question that asks participants to summarize their general decision-making approach, which would help uncover the characteristics depicted in this PDP.

The marginal effects of Extraversion and Neuroticism, Figure 3.4d and Figure 3.4f respectively, are fairly straightforward. Both increase in fairly linear fashions as their respective scores increase. Extraversion is largely related to engagement with others and has been found to be a positive predictor of cooperation in social dilemmas, so the results in Figure 3.4d are intuitive (Hirsh and Peterson, 2009; Lönnqvist et al., 2011; Ben-Ner and Kramer, 2011). Higher Neuroticism scores characterize participants who exhibit a greater degree of emotional instability or anxiety, especially when confronted with stressful situations (Kline et al., 2019). Therefore, it is also reasonable that higher Neuroticism scores are associated with a propensity to contribute more resources to ensure the disaster causes as little damage as possible.

The PDP for the potential savings of the participant, Figure 3.4e, predicts contributions that are closer to the "Fair Share" strategy at higher deciles, while that of the potential savings of the group, Figure 3.4m, reflects little to no effect and greater uncertainty. These trends are reversed in PDPs corresponding to the "Equal Share" model. In other words, the model that focuses on predicting the more individual-based strategy is affected more by an individual-based predictor for savings, and vice-versa for the more communal-based strategy.

Figures 3.4g and 3.4l show that higher values for other community members' resources predict larger, positive deviations from the "Fair Share" strategy. This dynamic stems from the fact that the other community members (i.e., the computer players) contribute fewer resources to reach the given Pareto efficient strategy when participants are over-contributing to a large degree and preemptively building up the shield.

The PDP for Agreeableness, Figure 3.4h, is counter-intuitive to expectations, with contributions trending lower and closer toward the Pareto efficient strategy at higher deciles. It should be noted that Agreeableness also has a negative effect in the MLRs (Table 2), and the PDPs corresponding to BART models that predict raw contributions follow similar trends as those depicted here (i.e., higher agreeableness predicts fewer contributions in general). Agreeableness is characterized as an orientation to the needs of others and is thought to be the predominant trait that underlies pro-social motivations (Zhao and Smillie, 2015). In fact, economic studies find it to be the trait most predictive of pro-social behaviors in cooperative dilemmas (Kline et al., 2019; Zhao and Smillie, 2015; Kagel and McGee, 2014; Pothos et al., 2011; Hilbig et al., 2013; Koole et al., 2001; Volk et al., 2011, 2012). Therefore, one would assume that higher Agreeableness scores, controlling for the effects of other variables, would predict increases in contribution amounts in our game. However, to the best of our knowledge, no other study that uses Big Five traits as predictors for economic dilemmas frame their game in a threat-avoidance setting. This distinction could be crucial here because the biological basis for Agreeableness can be mapped to the *CARE* circuit in humans, a primary emotional system identified by Jaak Panksepp that is aroused when animals feel the need to nurture or protect their familial unit (Montag and Panksepp, 2017; Montag and Davis, 2018; Marengo et al., 2021). Perhaps, then, the higher Agreeableness scores are predicting a heightened sense of urgency among participants to not squander resources when faced with the threat of an impending disaster in order to help ensure the survival of their familial unit (i.e., mamma bear is being protective of her cubs when threatened). This hypothesis is somewhat substantiated by the fact that Anderson (2010) found Agreeableness to be the only Big Five trait that predicts changes in preferences between Prospect Theory's loss and gain framing. However, more research needs to be done to properly investigate the role this trait has on predicting decisions in this type of CPR setting.

The PDPs for the community's starting shield, Figure 3.4i, and that of a participant's pre-game contributions, Figure 3.4j, have similar characteristics and underlying dynamics. Initial contributions in repeat social dilemmas often reflect the degree of cooperation exhibited by participants throughout the remainder of the game (Wunder et al., 2013; Nay and Vorobeychik, 2016). As such, higher pre-game contributions and starting shield values reflect a general propensity for participants to contribute more to the shield relative to the given Pareto efficient strategy.

The PDP for Conscientiousness, Figure 3.4k, is fairly intuitive. Conscientiousness reflects an attention to detail or ability to follow instructions (Kline et al., 2019; Zhao and Smillie, 2015). Thus, participants with higher scores for Conscientiousness are more likely to pay attention to the details of the game's dynamics and in turn make a Pareto efficient decision.

Lastly, the PDP for expected damage to the participant, Figure 3.4n slopes slightly upward at higher deciles, likely as a consequence of participants over-contributing to a greater extent when they feel more acutely threatened by the disaster. The opposite trend appears in the PDP for the expected damage to the group, Figure 3.4o, likely as a result of participants' general tendencies to over-contribute being magnified when overall damage is small. Unlike the PDPs for individual and group savings, these two remain the same across BART models.

### 3.4 Discussion

The BART models are able to predict participant deviations from the "Pareto Efficient: Fair Share" and "Pareto Efficient: Equal Share" strategies in our game much more effectively than the MLR counterparts. This fact is unsurprising as BART models belong to a class of powerful, non-parametric machine learning methods (tree-ensembles) that are better equipped to model complex and nonlinear relationships between variables than parametric regression (Chipman et al., 2010; Kapelner and Bleich, 2016). Additionally, these non-parametric procedures, often thought to be "black-boxes," become much more interpretable when combined with PDPs.

The PDPs are able to depict the marginal effect variables have on responses while properly controlling for other factors and without a-priori assuming the type of relationship that exists between inputs and responses. As illustrated in Figure 3.4b, the PDP for trait Openness depicts a nonlinear relationship and reveals that differentiating this trait into its two distinct but correlated *aspects* (i.e., intellect and openness proper) may be warranted in this setting (DeYoung et al., 2007). This finding could easily be overlooked in a typical linear regression analysis. As such, it might be prudent to utilize more advanced statistical learning techniques to reexamine assumptions of linearity in previous studies attempting to relate individuals' differences in attitudes and dispositions to decision-making in cooperative dilemmas.

Responses in our community resilience CPR game show that participants tend to over-contribute relative to the two aforementioned Pareto efficient strategies, and this finding may have ties to decision theory. Generally, individuals are risk-averse with respect to gains, and risk-aversion has been found to negatively correlate with contributions in traditional PGGs, which are framed in terms of gaining value via collective efforts (Levati et al., 2009; Teyssier, 2012; Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). Perhaps, the opposite occurs in cooperative dilemmas such as ours where the collective effort is framed in terms of avoiding detrimental outcomes. Related, our game also conveys to participants that they can effectively mitigate the consequences of the simulated disasters, so Protection Motivation Theory suggests that decision-makers will be more inclined to try to do so (Maddux and Rogers, 1983). More research is needed to test these ideas and similarly aid in recent efforts to unite aspects of psychometrics with formal decision theory (Bortoli et al., 2019; Boyce et al., 2016; Rustichini et al., 2016; Schröder and Gilboa Freedman, 2020).

Related, to help develop a better empirical understanding of how decision-makers respond in CPR settings where contributions are framed in terms of investments that bolster community resilience, we plan to systematically vary and study the effects of several game-level treatments in our study, including but not limited to the following factors: community size, multiplication factor, game length, and distribution of starting resources. Additionally, we plan to examine the effects of varying computer algorithms and/or comparing responses to different strategies. These strategies do not necessarily have to result in Pareto efficient outcomes either; they could be informed by participants goals stated before the game and/or even inferred by their Big Five profile. A cooperative dilemma can be thought of as a problem of individuals attempting to communicate different utility functions to one another. As such, by comparing responses to predefined strategies, we can assess how "well" participants played the game based on different criteria and/or identify which factors may contribute to deviations from their specified goals.

### 3.5 Conclusion

Community resilience plays a significant role in mitigating the risks associated with natural disasters. Public infrastructure that increases community resilience can often be conceptualized as a public good or common-pool resource. However, despite a growing body of game-theoretic research pertaining to disaster resilience, most studies in this context rely on simplifying behavioral assumptions and/or lack empirical justification for their employed decision models. Here, we design a web-based CPR game to examine patterns in participants' decisions to invest in public infrastructure that helps protect their communities from various disaster scenarios and use statistical learning techniques to construct empirically-based decision models from their responses.

In particular, we use BART models and MLRs to predict participant deviations from two different Pareto efficient strategies that constitute everyday notions of pro-social behavior. Participants' Big Five traits and in-game circumstances serve as input variables. We find that the BART models are able to accurately predict deviations from these strategies and perform much better than their MLR counterparts. Additionally, we find that trait Openness is a significant predictor across all models. However, PDPs derived from the BART models reveal that the relationship between Openness and participants' decisions is nonlinear. This finding suggests further differentiating this trait into its two distinct but correlated *aspects* (i.e., intellect and openness proper) may be warranted in this application and that in general, it might be prudent to utilize more advanced statistical learning approaches to reexamine assumptions of linearity in other studies that have related individuals' differences in temperaments to responses in cooperative dilemmas.

### 3.6 Supplementary Materials

All data and source code required to reproduce this analysis can be found on Open Science Framework: https://osf.io/6qruj/?view_only=bb30fb4f582d4a69baafc1ad242877e5

## CHAPTER 4

## An agent-based model for simulating economic disruptions due to floods along the inland waterways: A case study on the Port of Cates Landing

The U.S. inland waterways play a vital role in the domestic economy, but extreme weather events, especially floods, threaten to disrupt their operations. Here, we develop an agent-based model to simulate how businesses along the Upper Mississippi River may reasonably reroute shipments in response to empirically-based flood scenarios and calculate the subsequent economic impacts of such decisions using a multiregional inoperablility input-output model. We calibrate our simulation using approximate Bayesian computation with sequential Monte Carlo sampling, which in turn allows us highlight scenarios where a flood-resilient port, the Port of Cates Landing, is able to handle some of the rerouted shipments. We find that Illinois, Louisiana, Minnesota, and Missouri are the states that suffer the most production losses from waterway disruptions along the Upper Mississippi River and that agriculture and chemical manufacturing are the most impacted sectors. However, the Port of Cates landing is able to mitigate production losses, and its development costs, \$53M, are equivalent to the amount of savings realized during a "30-year" flood.

### 4.1 Introduction

The U.S. inland waterway system is an extensive collection of navigable channels, ports, and locks and dams that are crucial for maintaining the success of the domestic economy (MacKenzie et al., 2012; Camp et al., 2013). Every year, more than \$100 billion (2.2. billion tons) of commodities are transported along the inland waterways, which amounts to roughly 15% of total U.S. freight (Folga et al., 2009). These shipments have exemplary safety records, low carbon emissions, and low costs compared to other modes of transport (Philip and Johnson, 2018; Schweighofer, 2014). Perhaps most important though, ports along the inland waterways serve as vital hubs connecting a broader, multimodal network of barge, rail, and truck transport (MacKenzie et al., 2012).

Extreme weather events, especially floods, are among the primary risks that threaten to disrupt operations along the inland waterway system (Pregnolato et al., 2017). Extensive disruptions from such events can be devastating to the economic wellbeing of the affected region and nation (Folga et al., 2009). For example, in 2019, sections of the Upper Mississippi River (UMR) were closed for more than a month due to major flooding, which resulted in approximately \$1.2 billion of grain not being shipped (Fahie, 2019).

However, despite these known risks, researchers have devoted little attention to analyzing the effects of disruptions along the inland waterways, especially when compared to the amount of related work done on other modes of transport (Folga et al., 2009; MacKenzie et al., 2012). Among the few who have are Folga et al. (2009), who prescribed a systems-level approach to analyzing inland waterway disruptions where infrastructure capabilities, rerouting decisions, and economic impacts are all considered. In turn, MacKenzie et al. (2012) constructed an agent-based model (ABM) to investigate rerouting decisions of various industries in response to extended closures of the Port of Catoosa and linked their simulation to a multiregional economic model in order to quantify the impacts of these decisions. Similarly, Darayi et al. (2019) integrated commodity network flows with an economic interdependency model to analyze vulnerabilities to disruptions in Oklahoma's multimodal transportation network, which includes the inland waterways.

When a disruption occurs, freight scheduled to ship through an inland waterway port may be rerouted. In the case of a localized disruption (i.e., a single port is closed), shipments could theoretically be redirected through another nearby port if circumstances permit (MacKenzie et al., 2012). However, if the disruption is more widespread and an entire section of a river is closed, as is often the case with extreme weather events, alternative waterway routes are generally limited (Folga et al., 2009). As such, affected businesses may elect to seek out different modes of transport, namely truck or rail, or decide to leave their product on a barge or at port for the duration of the disruption (MacKenzie et al., 2012). These rerouting decisions are case-specific with regards to the costs, preferences, and constraints of each business making them, but they also have far-reaching consequences for the rest of the economy (MacKenzie et al., 2012).

Even though businesses may not consider the downstream impacts of their decisions of whether or not to reroute shipments, the macroeconomic inoperability input-output model (IIM) can evaluate how demand-side perturbations in given industry sectors ripple through the rest of an economy (Santos and Haimes, 2004; Haimes et al., 2005; Santos, 2006; MacKenzie et al., 2012). The IIM is a risk-based extension to the input-output (I-O) model, which is an established approach for estimating interdependencies between difference sectors of an economy (Leontief, 1936). Furthermore, the IIM can be extended via the multiregional inoperablility input-output model (MRIIM) to better capture differences in impacts among geographical sub-regions of an economy (Crowther and Haimes, 2009). Darayi et al. (2019) utilized the IIM to calculate economic impacts of disruptions along Oklahoma's multimodal transportation network, and MacKenzie et al. (2012) used a dynamic extension of the MRIIM in their ABM when estimating impacts of extended closures to the Port of Catoosa.

However, to the best of our knowledge, no systems-level risk assessment of inland waterway disruptions includes empirically grounded assumptions for the likelihood of disruptions occurring. Understanding the likelihood of events is a critical component in characterizing the risks of those events (Woodruff, 2005; Aven, 2016). Here, we extend existing approaches to analyzing disruption risks along the inland waterways by including a data-driven framework for modeling the likelihood of river closures occurring due to floods.

In particular, we build off the ABM developed by MacKenzie et al. (2012) that models how businesses may reasonably reroute shipments in response to port closures along the inland waterways and calculates the subsequent economic impacts of these decisions. By extending MacKenzie et al. (2012)'s model to factor in the likelihood of various flood scenarios, we can better quantify the economic risks associated with disruptions from floods.

Additionally, there is currently no guidance on how to parameterize existing waterway disruption models with data from extreme weather events. In general, calibrating models pertaining to disaster management is difficult because observational data is lacking for extreme weather events (Mahmassani, 1990; He and Liu, 2012; Jackson et al., 2016). Furthermore, typical statistical procedures for estimating model parameters, such as maximum likelihood estimation and general Bayesian inference, are often incompatible with simulations featuring complex and non-linear systems (Thiele et al., 2014; Fasiolo et al., 2016). Here, we show how to parametrize our ABM with inland waterway commodity flow data using a technique called approximate Bayesian computation with sequential Monte Carlo (ABC SMC). This analysis serves an example from which other researchers can draw insights regarding how to parameterize transportation disruption models.

We demonstrate our approaches with a case study along the UMR. The UMR is a key segment of the U.S. inland waterways, and every year more than 85 different commodities totalling 119 million tons are transported through this section of the Mississippi River (Fahie, 2019). However, the UMR region is also prone to flooding, recently exemplified by the historic floods in 2019 where water levels breached the "100-year" floodplains along several sections of the river (Associated Press, 2019; Gasparini and Yuill, 2020). Although we focus on the UMR in our analysis, our approach can be readily applied to other segments of the inland waterways.

With the empirically-based and calibrated model, we can reliably simulate regional production losses resulting from various flood disruption scenarios along the UMR. Moreover, we can analyze cases where the Port of Cates Landing, a public port located whose operations lie above the "500-year" flood plain near the mouth of the UMR, can handle some of the rerouted shipments and mitigate production losses for the region (Swaggart, 2021). This type of analysis serves as an example for how policy-makers can better assess when/where the development of public waterborne infrastructure may be warranted from an economic resilience standpoint.

### 4.2 Methods

Our methodology consists of three main tasks. First, we construct empirical distributions of the expected port closures along the UMR due to "XX-year" flood events; here, "XX-year" refers to the return period a flood (e.g., a "100-year" flood, or a flood expected to occur once every 100 years). These distributions are based on historical river gauge data and high-water operating procedures outlined in the Waterways Action Plan (WAP) (USCG, 2020; USACE, 2021). Second, we develop an ABM that simulates business' rerouting decisions in response to the flood scenarios and calculates the subsequent economic impacts of these decisions. Third, we calibrate model parameters with commodity flow data using a statistical technique called ABC SMC.

### 4.2.1 Flood Scenarios

In this section, we discuss how we model the expected port closures along the UMR per flood event return period. This task involves fitting Gumbel distributions to historical river gauge counts that exceed high-water operating thresholds as outlined in the WAP. We then combine these gauge-level forecasts into segment-level flood scenarios for the UMR.

The USACE (2021) maintains a database of readings from river gauges all across the U.S. Part of this data includes historical river stage levels for 25 gauges along the UMR, shown in Figure 4.1. For most of these gauges, daily stage level data are available dating back to 2014. As such, we use all available records from January 1, 2014 to December 31, 2020 in our analysis. It should be noted that this time frame is somewhat limited, especially when considering low-probability, high-consequence events, so we try our best to mention any related limitations of our analysis when appropriate throughout this paper.

There are several data processing steps required for our analysis. First, the ten northernmost gauges have a number of missing values and/or changes in the way stage level is measured during the first two years (see, Supplementary Materials). As such, we remove these records from our dataset and impute plausible replacement values via Multivariate Imputation by Chained Equations, using the **mice** package in *R*. These northernmost gauges rarely exceed flood thresholds (e.g., they were still well below closure levels during the 2019 floods), so the imputed values have little effect on the overall flood scenarios (USCG, 2020). Additionally, there are intermittent, missing daily entries throughout the rest of the records ($< 1\%$ of the remaining data). Since the daily river gauge data exhibit a high degree of temporal autocorrelation, we use time series interpolation to fill in these missing values.

The WAP delineates stage levels at which sections of the UMR should be closed due to high-water conditions (USCG, 2020). The WAP is a living document maintained by private stakeholders, the U.S. Coast Guard (USCG), and the U.S. Army Core of Engineers (USACE) that comprehensively describes inland waterway operating procedures in response to adverse weather conditions and is a key reason why the inland waterways have such an impeccable safety record (Philip and Johnson, 2018). Thus, using the historical river gauge data in conjunction with WAP operating procedures, we tabulate the number of days per year that each river gauge exceeds its corresponding high-water threshold. In other words, we create a dataset consisting of 25 rows (river gauges) and seven columns (years 2014 to 2020) of the number of days each gauge exceeded

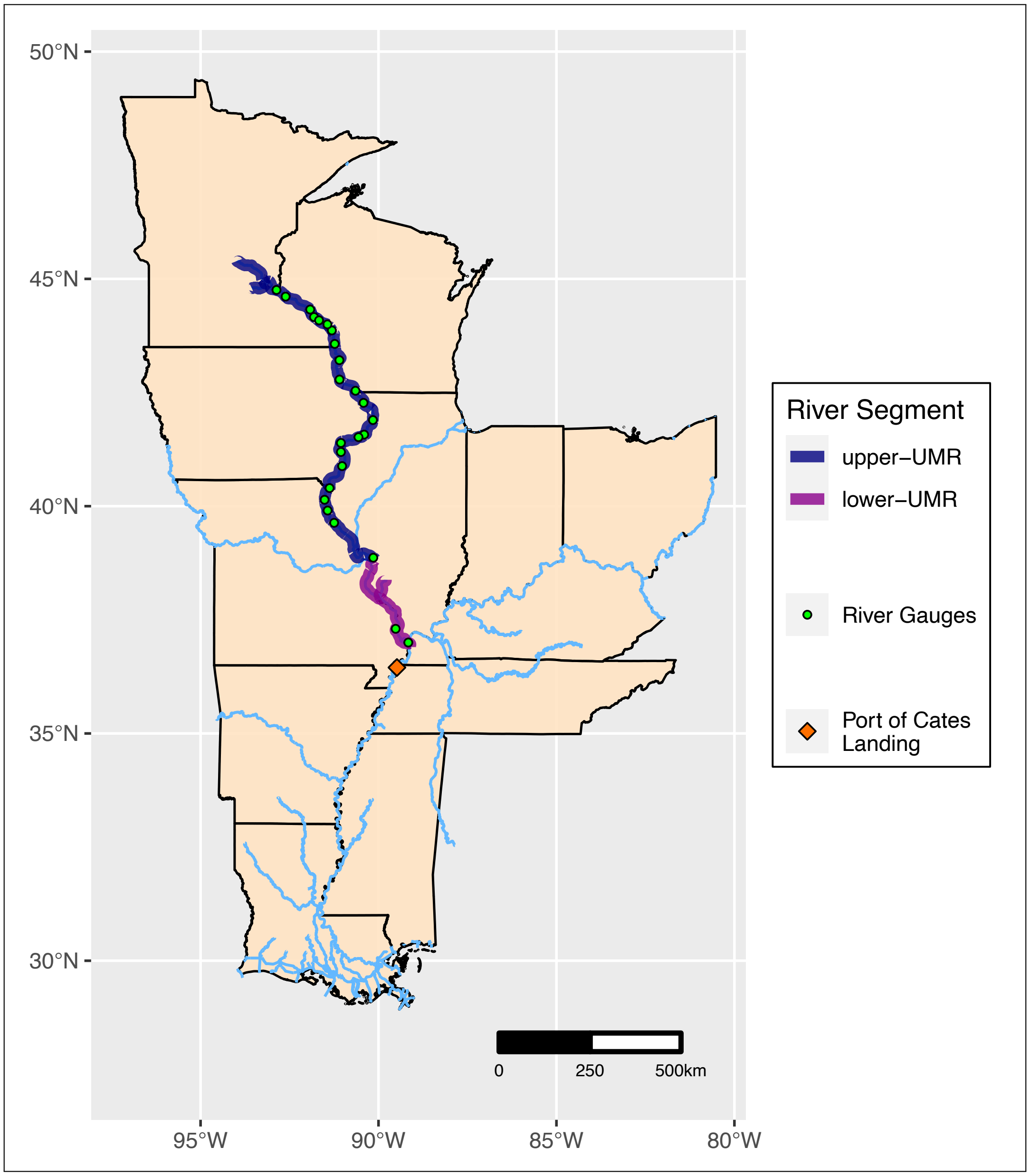



Figure 4.1: UMR Inland Waterway System

its flood threshold as defined by WAP protocols.

We then fit Gumbel distributions to each gauge in the dataset to forecast the return periods (measured in years) of the expected number of closures due to flood conditions. The Gumbel distribution is a special case of the generalized extreme value distribution, which models the asymptotic behavior of extreme values expressed as a Poisson process over a given number of samples (Gilleland and Katz, 2016). Researchers often use the Gumbel distribution to analyze flood data and in particular to model the return period of high-water levels (Bhagat, 2017; Solomon et al., 2013; Onen and Bagatur, 2017). The Gumbel distribution's density function is given below in Equation 4.1

$$G(z) = \exp\left[-\exp\left\{-\left(\frac{z-u}{\sigma}\right)\right\}\right], -\infty < z < \infty \tag{4.1}$$

Quantiles of this density function are of particular interest because they can be interpreted as return periods (Gilleland and Katz, 2016). In other words, we can expect a value corresponding to the $q^{th}$ quantile to be exceeded every $1/(1-q)$ periods (e.g., the 95% quantile will be exceeded every $1/(1-0.95) = 20$ periods). Thus, by fitting Gumbel distributions to number of days river gauges exceed high-water thresholds per year, we are able to generate forecasts for the expected number of days sections of the river are closed due to a "XX-year" flood event.

Note, flood return periods are typically described in terms of maximum flood levels. However, maximum water levels and counts of days exceeding high-water thresholds are highly correlated and can both be characterized as Poisson processes (Gilleland and Katz, 2016). As such, from a conceptual standpoint, we do not see describing return periods in terms of counts of flood closures, instead of return levels, as a major issue. Additionally, from a practical standpoint, we validate our forecasts against river closures resulting from a "100-year" flood and found that other distributions commonly used to fit extreme value data, namely the Log-Normal, Weibull, and General Pareto distributions, did not perform as well as the Gumbel distribution and/or had a harder time estimating model parameters. Onen and Bagatur (2017) also noted that Gumbel distributions tend to be more effective than other extreme value distributions when dealing with fairly small sample sizes, which we have in our case with only seven annual data points per gauge.

Next, we combine each gauge's distribution into two river segment forecasts, as shown in Figure 4.1 and Table 4.1. The first forecast consists of average predictions from river gauges that span the upper portion of the UMR (i.e., Hastings Lock and Dam to Saverton Lock and Dam). The second forecast consists of average predictions from gauges of the lower portion of the UMR (i.e., Melvin Price Lock and Dam to Cairo Point). Table 4.1 shows predictions for the expected number of days each segment of the UMR is closed due to "10-year", "20-year", "50-year", "100-year", and "500-year" floods.

Table 4.1: UMR Expected Days Closed per Flood Return Period

| | 10yr | 20yr | 50yr | 100yr | 500yr |
|---|---|---|---|---|---|
| upper-UMR | 12.1 | 15.4 | 19.7 | 22.9 | 30.3 |
| lower-UMR | 23.6 | 30.0 | 38.3 | 44.6 | 59.1 |

We aggregated predictions from each river gauge into these two segment-level forecasts for several reasons. First, averaging predictions helps diminish the effects from noise in individual gauge forecasts, which is more of an issue when we only have seven data points (years) from which to fit distributions. Second, data from gauges within each of these segments exhibit similar characteristics, likely due to geographical factors: the upper-UMR does not have any major tributaries, whereas both the Illinois River and Missouri River flow into the lower-UMR, thereby compounding impacts from floods (see, Figure 4.1). As such, this natural geographical distinction serves as an intuitive divide in the UMR. Lastly, segment-level forecasts tie much more closely to the logic embedded in our ABM (discussed later).

We use documented river closures during the UMR floods of 2019, which breached "100-year" flood plains along several parts of the UMR, as reference points to validate our forecasts. During the 2019 UMR floods, several sections of the upper-UMR were closed for eighteen days, and the St. Louis Harbour, the major port along the lower-UMR, was closed for 51 days (Fahie, 2019). These data points are close to corresponding predictions in Table 4.1, confirming that our forecasted river closures for a "100-year" flood scenario are reasonable. As more river gauge data become available in the future, these forecasts will improve, and ideally, we'll also be able to validate different return periods. Additionally, we found that business decisions and resulting economic impacts in our ABM are insensitive to small deviations in the number of days each segment of the river is closed for floods whose return periods exceed five years (see, Supplementary Materials).

### 4.2.2 Agent-based Model

Our ABM features a decision model where various companies determine how to reroute their shipments in response to disruptions from flooding along the inland waterways. The span of the disruption is based on the return periods of the previously described flood scenarios. Company decisions are then linked to a MRIIM to determine overall economic impacts for the region. Our simulation methodology is largely based on the work done by MacKenzie et al. (2012). Below, we use ODD + D protocol, an extension to the standard ODD protocol, to describe our ABM (Grimm et al., 2006; Müller et al., 2013).

#### 4.2.2.1 Overview

**Purpose**

The purpose of the simulation is to better understand how inland waterway disruptions due to flooding along the UMR affect the regional economy. Once calibrated, the model is also used to highlight cases in which the Port of Cates Landing, a public port whose operations lie above the "500-year" flood plain, can handle some of the rerouted shipments and help mitigate production losses. Our model is primarily designed for academic researchers interested in analyzing risks of disruptions along the inland waterways, but policy-makers may also find utility in its applications and insights.

**Entities, state variables and scales**

The simulation consists of two entity types. The first is a flood, which has two attributes: affected river segment (i.e., upper-UMR or lower-UMR, as depicted in Figure 4.1) and return period, as described in the previous section. From these two attributes, we forecast the number of days each segment of the river is expected to be closed. The second type of entity is a company/business that had planned to ship its product through the UMR (MacKenzie et al., 2012). Each company is characterized by its state of operation, industry sector, and shipment destination. The featured region includes the twelve states that have significant amounts of inbound/outbound shipments through the UMR, as revealed by USACE's Waterborne Commerce Statistics Center (WCSC) (USACE, 2017). These states are Minnesota, Wisconsin, Iowa, Missouri, Illinois, Indiana, Ohio, Kentucky, Tennessee, Arkansas, Mississippi, and Louisiana (dark green in Figure 4.1). Industry sector is defined by the 2017 North American Industry Classification System (NAICS), which consists of 71 different industry codes (US BEA, 2017).

Similar to MacKenzie et al. (2012), we assume there is one-to-one correspondence between a given company and the unique combination of a state's industrial sector and its shipment destination (i.e., 12 origin states $\times$ 71 NAICS sectors $\times$ 12 destination states = 10,224 possible company agents). However, as MacKenzie et al. (2012) also noted in their case, most of these companies do not have commodity flows associated with them; inland waterway transports predominantly consist of a few types of products/sectors, some of the main ones being agriculture, coal, petroleum, chemicals, and mining. As such, only 88 of the company agents in our simulation have non-zero commodity flows, and thus, these 88 companies comprise the set of active agents in our simulation (see, Supplementary Materials for details). MacKenzie et al. (2012) found that increasing the number of companies per origin-sector-destination combination did not affect outcomes because the rerouting logic is applied in the same manner across companies. Furthermore, we only have state-level commodity flow data available, so constructing more granular company agents is not possible without making tenuous assumptions.

Spatial aspects of the simulation are modeled explicitly via the location of the flood disruptions (i.e., affected river segments) and by companies' shipping routes. Flood disruptions occur along the upper-UMR and lower-UMR as shown in Figure 4.1 and Table 4.1. The spatial elements associated with shipping routes is explained in more detail in the following paragraphs.

**Process overview and scheduling**

Each simulation run is modeled as one year. Prior to each run, the model user specifies a list of flood return periods he or she wishes to analyze. The simulation then determines the expected number of days each segment of the UMR is expected to be closed based on the flood return period(s). Then, for each day the river is closed, the companies decide whether or not to reroute their product via an alternative shipping method or leave their product at port based on several factors, discussed in detail in the following paragraphs. In turn, a MRIIM calculates the cascading economic impacts of the companies' decisions.

#### 4.2.2.2 Design Concepts

**Theoretical and empirical background**

Our model is fundamentally based on the ABM developed by MacKenzie et al. (2012) that simulates how businesses may reasonably reroute shipments in response to port closures along the inland waterways and calculates the subsequent economic impacts of such decisions. Companies shipping products have two main objectives: minimize total shipping costs and satisfy on-time shipping requirements (MacKenzie et al., 2012). As such, their decisions in response to disruptions are determined by balancing a trade-off between the cost of shipping products via alternative routes for an on-time delivery ($C_{alt}$) versus the cost of keeping their product at port and failing to meet delivery requirements ($C_{port}$). If $C_{alt} < C_{port}$, the company will ship its products via an alternative method.

In the baseline version of our simulation, similar to MacKenzie et al. (2012), we assume that the only alternative shipping method considered by the companies is rail. Rail is cheaper than truck, and the UMR has no alternative inland waterway routes (MacKenzie et al., 2012). However, once we calibrate the baseline version of the model, we develop a second version of the model where a "500-year" flood resilient port, the Port of Cates Landing, is able to handle some of the rerouted shipments. Because the port operates above a "500-year" flood plain near the mouth of the UMR (start of the Lower Mississippi River, Figure 4.1), we reasonably assume that shipments originating along the UMR and traveling southward below the port will be able to reroute via rail past the flooded UMR to the port where they can make the rest of the journey along the waterways. Historical river gauge data and WAP protocols reveal that the Lower Mississippi River has a much lower risk of being closed due to high-water conditions than the UMR (USCG, 2020; USACE, 2021).

Companies' routing decisions are propagated through a MRIIM, so we can estimate the impact these decisions have on regional economic production. If companies ship their product via an alternative route when a disruption occurs, we assume their product reaches customers in a timely manner (i.e., demand from intermediate industries is satisfied) (MacKenzie et al., 2012). However, if companies decide to keep their product at port until the UMR opens, this delay leads to other, interdependent industries experiencing disruptions in production (MacKenzie et al., 2012). We measure these disruptions with the MRIIM.

The MRIIM framework is based on Leontief's input-output (I-O) model that describes the interdependent nature of sectors of an economy (Leontief, 1936). The inoperability input-ouput model (IIM) conceptualizes the I-O framework in terms of degraded production levels relative to baseline conditions and can be used to estimate how demand-side perturbations in given sectors affect the rest of an economy (Santos and Haimes, 2004; Haimes et al., 2005; Santos, 2006). The MRIIM, or multi-regional IIM, extends the IIM to calibrate impacts to multiple, often more granular locales and is specified by Equation 4.2 (Santos and Haimes, 2004; Crowther and Haimes, 2009):

$$\begin{pmatrix} \tilde{\mathbf{q}}^1 \\ \tilde{\mathbf{q}}^2 \\ \vdots \\ \tilde{\mathbf{q}}^p \end{pmatrix} = \mathbf{T}^* \begin{pmatrix} \mathbf{A}^{*1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^{*2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{A}^{*p} \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{q}}^1 \\ \tilde{\mathbf{q}}^2 \\ \vdots \\ \tilde{\mathbf{q}}^p \end{pmatrix} + \mathbf{T}^* \begin{pmatrix} \tilde{\mathbf{c}}^{*1} \\ \tilde{\mathbf{c}}^{*2} \\ \vdots \\ \tilde{\mathbf{c}}^{*p} \end{pmatrix} \tag{4.2}$$

where

$\tilde{\mathbf{q}}^r$ = an inoperability vector of length $n$ consisting of the difference between normal production levels and disrupted production levels, expressed as a percentage of normal production levels, of the $n^{th}$ industry sector in region $r$ of $p$ total regions;

$\mathbf{T}^*$ = $[\text{diag}(\tilde{\mathbf{x}}^1, \tilde{\mathbf{x}}^2, \dots, \tilde{\mathbf{x}}^p)]^{-1}\mathbf{T}[\text{diag}(\tilde{\mathbf{x}}^1, \tilde{\mathbf{x}}^2, \dots, \tilde{\mathbf{x}}^p)]$;

$\tilde{\mathbf{x}}^r$ = a vector of length $n$ consisting of industry sector production in region $r$;

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}^{11} & \mathbf{T}^{12} & \dots & \mathbf{T}^{1p} \\ \mathbf{T}^{21} & \mathbf{T}^{22} & \dots & \mathbf{T}^{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{T}^{p1} & \mathbf{T}^{p2} & \dots & \mathbf{T}^{pp} \end{pmatrix};$$

$\mathbf{T}^{rs}$ = an $n \times n$ trade interdependency matrix consisting of the proportion of a commodity consumed in region $s$ that is produced in region $r$;

$\mathbf{A}^{*r}$ = the inoperability matrix for region $r$, $[\text{diag}(\tilde{\mathbf{x}}^r)]^{-1}\mathbf{A}^r[\text{diag}(\tilde{\mathbf{x}}^r)]$;

$\mathbf{A}^r$ = an $n \times n$ industry interdependency matrix of region $r$ composed of elements $a^r_{ij}$;

$$a^r_{ij} = \begin{cases} l^r_i a_{ij} & , l^r_i < 1 \\ a_{ij} & , l^r_i \geq 1 \end{cases};$$

$a_{ij}$ = the input of industry sector $i$ to $j$, expressed as a proportion of the total production inputs to industry sector $j$;

$l^r_i$ = the location quotient, $\frac{x^r_i/x^r}{x_i/x}$;

$x^r_i$ = industry sector $i$'s production in region $r$;

$x^r$ = total economic production in region $r$;

$x_i$ $=$ industry sector $i$'s production across the nation;

$x$ $=$ total national economic production;

$\tilde{\mathbf{c}}^{*r}$ $=$ a demand-side perturbation vector of length $n$ consisting of the difference between normal demand and disrupted demand, expressed as a percentage of normal production levels, of the $n^{th}$ industry sector in region $r$ of $p$ total regions

In our simulation, there are 71 NAICS sectors (i.e., $n = 71$). Additionally, the aforementioned twelve states that have significant amounts of inbound/outbound shipments through the UMR comprise the twelve regions (i.e., $p = 12$ states, and $r$ represents a given state). Company shipments that do not reach their destination in a timely manner are modeled as perturbations in demand (Horowitz and Planting, 2009; MacKenzie et al., 2012). In other words, if companies in state $r$ decide to keep their products at port, the value of these products, expressed as a proportion of that state-sector's normal shipments, forms the corresponding entry in $\tilde{\mathbf{c}}^{*r}$. Given $\tilde{\mathbf{c}}^{*r}$ for all twelve states, we then solve for $\vec{\mathbf{q}}$ (vector of length $n \times p$) in Equation 4.2 to estimate the disruptions to each industry sector in each state.

### Individual decision-making

Decision-making is modelled at the company level. As mentioned, when faced with a disruption, businesses that ship their products along the UMR must make a trade-off between paying extra transportation costs to ship via alternative methods for an on-time delivery ($C_{alt}$) or keeping their product at port and failing to meet delivery requirements ($C_{port}$).

$C_{alt}$ includes the base cost of companies shipping via an alternative method multiplied by a discount rate $(1-\beta)$, representing a company's desire to make an on-time delivery, and is specified by Equation 4.3 (MacKenzie et al., 2012). For equivalent distances and weights, the base cost of shipping via waterways is less than that of rail, approximately 0.97 cents per ton-mile versus 2.53 cents per ton-mile respectively (MacKenzie et al., 2012; AOPOA, 2010). In the baseline version of the model, companies only consider rail as an alternative shipping method, so the ton-miles of waterway shipments would be zero in Equation 4.3. Similar to MacKenzie et al. (2012), we assume there are no capacity or availability constraints on the rail shipments.

$$C_{alt} = (RailTonMiles_{daily} \times 2.53 + WaterTonMiles_{daily} \times 0.97) \times (1-\beta) \tag{4.3}$$

However, in the second version of the model, we assume that some of the shipments originating along the UMR and traveling southward would be able to reroute via rail past the flooded river to the Port of Cates Landing, where they would then be transported via the waterways to their final destination. As such, in this version of the simulation, Equation 4.3 may include both rail and water transportation costs. We assume the

Port of Cates Landing has a throughput capacity of 1.6 million tons per year, approximately 4,384 tons per day, based on economic reports and industry insights (Arik and Penn, 2009). The port's capacity is allocated in proportion to demand for using the port (i.e., if a company finds the cost of rerouting their shipments through Cates Landing to be less than that of using rail exclusively or keeping their products at port, that company will have demand for using the port). Once the throughput capacity is reached, any additional shipments will have to be transported via rail exclusively or remain at their original port.

$C_{port}$ is calculated as a penalty cost, representing a late delivery fee imposed by the customer and/or the perishability of a product, plus the base cost of shipping products via the inland waterways, as seen in Equation 4.4 (MacKenzie et al., 2012). In our simulation, the penalty cost is modeled as a percentage ($\alpha$) of the daily value of the products at port scaled linearly by the number of days the product will remain at port (i.e., the length of the river closure) (MacKenzie et al., 2012).

$$C_{port} = (Value_{daily} \times \alpha \times Days_{closed}) + (WaterTonMiles_{daily} \times 0.97) \tag{4.4}$$

As mentioned, the number of days each segment of the UMR is closed is determined by the return period of the forecasted flood event. Any waterway shipment that would pass through closed segments of the UMR will be delayed for the maximum closure length of the segments it must traverse. For example, a "10-year flood" is expected to result in approximately twelve days of flood closures for the upper-UMR and 24 days for the lower-UMR (see, Table 4.1). As such, any waterway shipments from Minnesota to Iowa or Wisconsin would be blocked for twelve days, but any shipments from Minnesota to Illinois, Missouri, Louisiana, etc. would be blocked for 24 days.

Route distances are estimated as the geographical distances between centroids of origin and destination states; we use the centroid of states to approximate distances because we do not have port-level commodity flow data. However, when the Port of Cates Landing is used to reroute shipments in the second version of the model, this route distance is calculated as the geographical distance from the centroid of the origin state to the Port of Cates Landing to the centroid of the destination state. Metrics for daily ton-mile commodity flows are based on the 2017 state-level shipments from the Waterborne Commerce Statistics (USACE, 2017).

Our analysis implies that disruptions occur during the late spring, as that is often when flooding along the UMR occurs. However, we uniformly parse the annualized commodity flows into daily flows in our simulation because we do not have temporally detailed data to model differences in seasonal shipments. Similarly, we do not place any constraints on the availability barge transport (i.e., waterway shipments can be transported as long as the river is not closed and do not have to wait for available barges). Both of these assumptions can be expanded upon in the future should more detailed spatio-temporal waterborne commodity data become available.

**Learning**

The decision rules explained above remain constant across the duration of the simulation (i.e., companies do not learn and/or change their behavior based on the decisions of others).

**Individual sensing**

Companies are assumed to have perfect knowledge of the number of days the river segments will remain closed and of the remaining capacity of the Port of Cates Landing. These assumptions are simplifying ones that allow us demonstrate our overall methodological framework. However, in future works, we plan to implement a more sophisticated and ideally empirically grounded behavioral models that capture uncertainties of river closure forecasts and how these uncertainties may affect company responses.

**Individual prediction**

Companies are assumed to have perfect knowledge of the factors that influence their decisions, so there is no prediction involved with their decisions.

**Interaction**

Companies interact only indirectly with one another. The routing decisions companies make in response to the disruptions propagate through the regional economy, as determined by a MRIIM, which indirectly affect each other's losses in economic production. Additionally, in the second version of the simulation, companies indirectly compete for the availability of the Port of Cates Landing.

**Collectives**

Companies do not belong to or form any collectives in this simulation.

**Heterogeneity**

Companies are heterogeneous with respect to their individual attributes (i.e., location, sector, destination) but are homogeneous with respect to how they make decisions.

**Stochasticity**

We use ABC SMC to obtain a posterior distribution for $\alpha$ and $\beta$ (discussed later). This distribution serves as a stochastic input to our simulation.

**Observation**

Company-level data collected from our ABM include companies' routing decisions and the corresponding weights, values, and costs of their shipments. We use company-level metrics to make sure our model is operating as expected. However, the key outputs we use for calibration, reporting, and inference are the posterior distributions of the emergent state-level and sector-level production losses.

#### 4.2.2.3 Details

**Implementation details**

Our model is implemented in *R* programming language. The annotated source code and accompanying data files can be found in the Supplementary Materials.

**Initialization**

The model is initialized by the user specifying a list of flood return period(s) of interest and values for $\beta$ and $\alpha$. We initialize the model with "10-year", "20-year", "50-year", "100-year", and "500-year" return periods listed as flood scenarios and an approximated posterior distribution for $\beta$ and $\alpha$ that is derived from ABC SMC.

**Input data**

Data for company shipments are derived from the 2017 state-to-state inland waterway commodity flows included in the WCSC (USACE, 2017). Commodity flows in the WCSC data are categorized slightly differently from those in the NAICS, so we re-categorize the former into the latter.

Data for the MRIIM primarily come from the U.S. Bureau of Economic Analysis (BEA) (US BEA, 2017). The BEA "Make" and "Use" matrices delineate the amounts of different commodities produced and consumed by various industries respectively and are used to derive the Leontiff coeffecients, $a_{ij}$, in Equation 4.2 (Santos and Haimes, 2004). Similarly, BEA data for gross domestic products (GDP) are used to calculate state-industry production vectors and location quotients. Lastly, the interdependency trade matrix, $\mathbf{T}$, is based on data from U.S. Census Bureau's Commodity Flow Surveys (US Census Bureau, 2017; Isard, 1998).

**Submodels**

The simulation includes three submodels: 1) the flood scenario forecast, 2) the company decision model, and 3) the MRIIM. We continuously tested all of these submodels for congruence between model logic and outcomes as we developed the simulation. Additionally, we calibrate the baseline version of the simulation (i.e., where only rail is used as an alternative shipping method) with commodity flow data from the 2019

UMR floods via ABC SMC. All model parameters and submodel routines are documented in the annotated source code provided in the Supplementary Materials.

### 4.2.3 ABC SMC Model Calibration

We use a statistical technique called approximate Bayesian computation with sequential Monte Carlo (ABC SMC) to calibrate our baseline simulation and construct posterior estimates of $\beta$ (the company discount rate applied to on-time deliveries in Equation 4.3) and $\alpha$ (the company late penalty rate applied to delayed shipments in Equation 4.4). ABC SMC is a particular case of a broader statistical technique known as approximate Bayesian computation (ABC) (Tavaré et al., 1997; Pritchard et al., 1999).

Parameterizing ABMs is often difficult because complex simulations are generally incompatible with traditional Bayesian inference and/or maximum likelihood estimation because their likelihood functions are often unknown or intractable (Thiele et al., 2014; Fasiolo et al., 2016). As such, the premise of ABC is to approximate the evaluation of these intractable or unknown likelihood functions by sampling from various parameter values and comparing corresponding simulation outcomes to observed data; parameter values associated with model runs that are within a given threshold of the observed data are retained for analysis and form the approximations of the posterior distributions of those parameters (Thiele et al., 2014; Grazzini et al., 2017).

ABC SMC aims to achieve efficient sampling of parameter values while simultaneously allowing for small error thresholds when comparing simulation outcomes to observed data, which leads to more effective and accurate approximations of posterior distributions (Beaumont, 2019). The particular implementation of ABC SMC we use is based on sequential importance sampling and can be regarded as a version of population Monte Carlo (PMC) (Sisson et al., 2007; Beaumont et al., 2009; Toni et al., 2009). The overall function of ABC PMC is to successively fit proposal distributions that progressively do a better job of sampling from the approximated posterior distributions.

The exact algorithm we use is taken from Beaumont et al. (2009) and is specified in Algorithm 1 (Cisewski, 2014). The initial proposal density, or target distribution, is set to the prior distributions of the parameters of interest, $\pi(\theta)$. Draws from $\pi(\theta)$ are propagated through the simulation, $f(y \mid \theta_i^{(1)})$, until the specified kernel function, $\rho(y,x)$, describing distances between the simulated outcome(s), $y$, and observed value(s), $x$, is less than the initial error threshold, $\varepsilon_1$. This process repeats until $N$ particles (draws) meet this criteria, which are then given equal sampling weights, $w_i^{(1)}$. Both $\rho(y,x)$ and $\varepsilon_1$ are chosen by the user.

After the initial timestep, the error threshold decreases such that $\varepsilon_t \leq \varepsilon_{t-1}$, which is how the ABC PMC algorithm progressively samples better proposal distributions, and $\tau_t^2$ is set to twice the variance of the previous timestep's proposal distribution (Beaumont et al., 2009). Draws from the improved proposal distribution,

$\theta_i^*$, are made based on the particle weights from the previous timestep. However, before they are propagated through the forward model, they are first resampled from $\mathcal{N}(\theta_i^*, \tau_t^2)$ to avoid particle degeneracy. This process repeats until $N$ particles (draws) meet this criteria, which are then re-weighted in proportion to their respective probabilities of ending up at their current values, $w_i^{(t)}$. The algorithm repeats for a number of timesteps specified by the user.

---

**Algorithm 1:** ABC PMC

---

**if** *t = 1 (timesteps)* **then**
  **for** $i = 1, \ldots, N$ *(particles)* **do**
    Generate $\theta_i^{(1)} \sim \pi(\theta)$ and $x \sim f\left(y \mid \theta_i^{(1)}\right)$ until $\rho(y,x) \leq \varepsilon_1$ ;
    Set $w_i^{(1)} = N^{-1}$ ;
  **end**
**else if** $t = 2, \ldots, T$ **then**
  Set $\varepsilon_t \leq \varepsilon_{t-1}$ ;
  Set $\tau_t^2 = 2 \cdot var\left(\theta_{1:N}^{(t-1)}\right)$ ;
  **for** $i = 1, \ldots, N$ **do**
    Draw $\theta_i^* \sim multinomial\left(\theta_{1:N}^{(t-1)}, w_{1:N}^{(t-1)}\right)$ ;
    Generate $\left(\theta_i^{(t)} \mid \theta_i^*\right) \sim \mathcal{N}\left(\theta_i^*, \tau_t^2\right)$ and $x \sim f\left(y \mid \theta_i^{(t)}\right)$ until $\rho(y,x) < \varepsilon_t$ ;
    Set $w_i^{(t)} \propto \pi\left(\theta_i^{(t)}\right) / \sum_{j=1}^{N} w_j^{t-1} \phi\left(\tau_t^{-1}\left(\theta_i^{(t)} - \theta_j^{(t-1)}\right)\right)$ where $\phi(\cdot)$ is the density function $\mathcal{N}(0,1)$ ;
  **end**

---

We initialize the ABC PMC algorithm with $N = 100$ particles and $T = 15$ timesteps based on existing guidance and by exploring trade-offs between reductions in $\rho(y,x)$ and increases in computation time (Cisewski, 2014). For similar reasons, we set $\varepsilon_t$ to be less than or equal to the $75^{th}$ quantile of the error values in the previous timestep. We define $\rho(y,x)$ as the absolute difference between simulated and observed losses in agricultural shipments through the UMR. Fahie (2019) found that relative to 2017 commodity flows, approximately 7.4 million tons (\$1.2B) of agriculture were not shipped through the UMR due to the "100-year" flooding in 2019. Related, the Special Hazard Events and Losses Database for the United States (SHELDUS) determined that there was \$112M worth of crop damage in the UMR region due to the 2019 flooding (Center for Emergency Management and Homeland Security, 2018). Thus, the amount of agriculture not shipped net crop damages, \$1.1B, constitutes the observed value, $x$, on which we calibrate our company decision model (i.e., the presumed amount of agriculture that companies decided not to reroute). In turn, $y$ is the total value of agricultural products that companies do not ship in our simulation due to a "100-year" flood event. We use the baseline version of our simulation (i.e., the version where rail is the only alternative shipping method) for calibration because the Port of Cates Landing was not operational in 2019. Lastly, we initialize $\varepsilon_1$ to be \$500M in agricultural shipments (i.e., roughly a 50% error), but by the end of the ABC PMC algorithm, $\varepsilon_{15}$ is only \$50M (i.e., approximately a 5% error).

We place Beta distribution priors on both $\beta$ and $\alpha$ to help constrain the estimation of these rate parameters to an intuitive range of 0 to 1. Since there is no empirical justification for what $\pi(\beta)$ should be, we set a weak prior over the entire range (i.e., $\pi(\beta) \sim Beta(1.25, 1.25)$). Similarly, we only have anecdotal evidence for what $\pi(\alpha)$ might entail. In 2010, Wal-Mart imposed a flat 3% penalty on the cost of goods of shipments that were more than four days late, and Anjoran (2009) suggest a 3% late-fee for the first week and 10% for each additional week (Whalen and Painter, 2010; MacKenzie et al., 2012). MacKenzie et al. (2012) used Monte Carlo simulation to observe the effects of varying $\alpha$ in their simulation and found that changes in regional production were sensitive at lower values of $\alpha$ (i.e., from around 0 to 20%) but less pronounced at higher values. As such, we wanted $\pi(\alpha)$ to be fairly weak but also for it to skew slightly toward these lower values (i.e., $\pi(\alpha) \sim Beta(1.5, 3)$).

### 4.3 Results

Results of the ABC PMC algorithm are shown in Figure 4.2. As seen in Figure 4.2a, the ABC PMC algorithm quickly converges to a stable error. It is customary for ABC PMC algorithms to have exponential decreases in $\varepsilon_t$ in early timesteps and diminishing improvements over subsequent timesteps; the opposite tends to occur for computation time (Cisewski, 2014).

Figure 4.2b shows the joint and marginal posterior distributions for $\alpha$ and $\beta$ approximated by the ABC PMC algorithm at $t = 15$. As seen, estimates for $\alpha$ are contained to a fairly narrow band of values, from around 5% to 10%. Those for $\beta$ are much broader, ranging from 0% to 50%. Additionally, there is a strong, negative correlation between the two, which makes sense given they have interdependent and competing roles in Equations 3.2 and 3.3. It should be that we sample from the joint distribution $\alpha$ and $\beta$ in the ABC PMC algorithm and in our ABM precisely because of their interdependent nature. The posterior approximation depicted in Figure 4.2b is largely unchanged from approximations beginning around timestep four (see, Supplementary Materials), which is expected based on Figure 4.2a.

Figure 4.2c depicts a sensitivity analysis of $\alpha$ and $\beta$ with respect to the amount of agriculture not shipped in the baseline version of the simulation due to “100-year” flood (i.e., the circumstances and scenario on which we calibrate our model). As seen, there is monotonic relationship between these parameters and the amount of agriculture that does not ship, which is likely why the ABC PMC algorithm was able to converge so quickly to a stable approximation. As such, we have a high degree of confidence that the ABC PMC algorithm was able to sufficiently explore the parameter space and that our approximation is about as accurate as it can

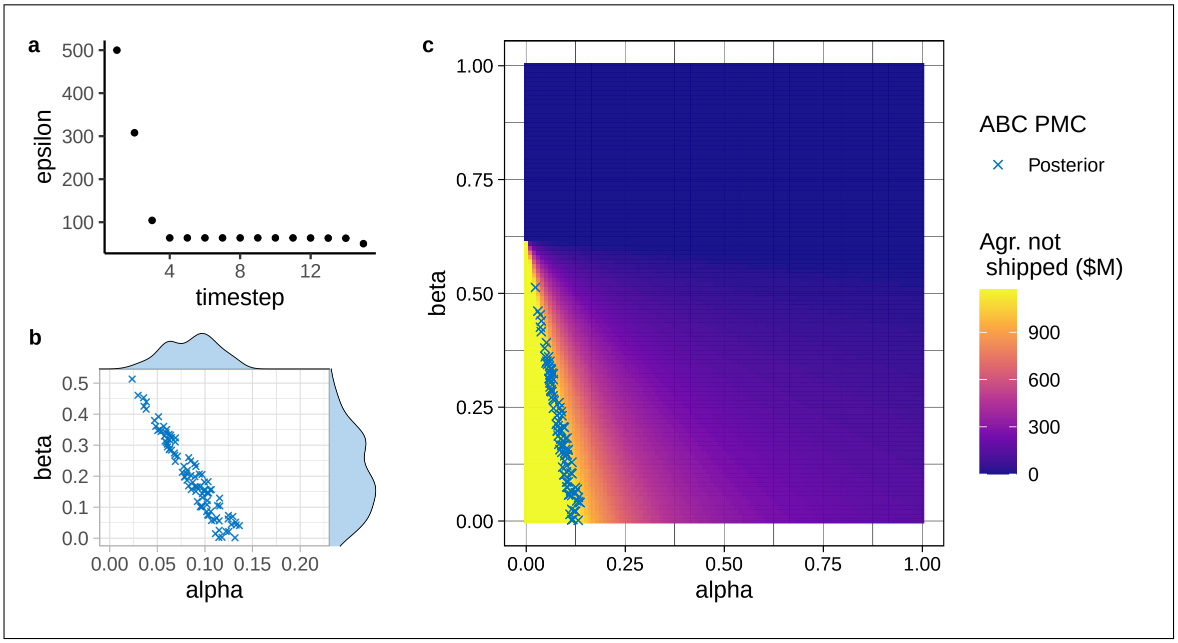



Figure 4.2: ABC PMC (a) $\varepsilon_t$ vs. $t$; (b) approximation of posterior at $t = 15$; (c) agriculture not shipped ($M) due to a "100-year" flood event for various $\alpha$ and $\beta$

feasibly be using this approach. Also noteworthy, $\alpha$ exhibits the same trend in being sensitive at lower values but less so at higher values that MacKenzie et al. (2012) observed.

Given that we now have a sufficient approximation of the posterior of $\alpha$ and $\beta$ (i.e., a calibrated model), we can properly simulate economic outcomes of various flood return periods. More specifically, we sample from the posterior of $\alpha$ and $\beta$ to obtain a posterior distribution of regional production losses, defined as $Q = \vec{\mathbf{q}}^T\vec{\mathbf{x}}$, as shown in Figure 4.3.

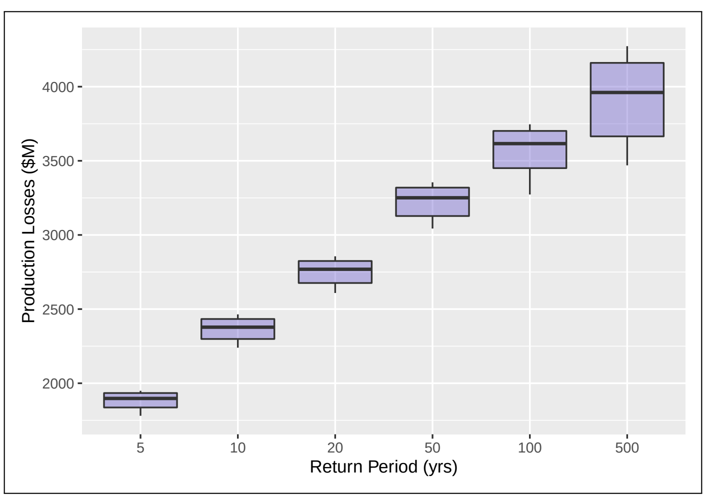



Figure 4.3: Posterior Regional Production Losses (baseline simulation)

As seen, more severe floods result in greater production losses, but there is a diminishing effect (e.g., a "10-year" flood leads to roughly 25% greater losses in production than a "5-year" flood, but a "500-year" flood results in only around 15% more than a "100-year" flood). This is due to the increase in the number of days the river is closed becoming too costly for companies to keep their products at port, so many of them end up shipping their products via rail, which in turn satisfies intermediate demand (i.e., no lost production).

The posterior production losses can also be broken down by state, $Q^r = [\tilde{\mathbf{q}}^r]^T \tilde{\mathbf{x}}^r$, as shown in Figure 4.4. The most impacted states are Illinois, Missouri, and Minnesota, all of which are located along the UMR and have significant amounts of outbound waterway shipments. Additionally, Louisiana suffers a lot of production losses because it is located at the mouth of the Mississippi River and receives much of the southbound shipments originating from the UMR (see, Figure 4.1).


Figure 4.4: Posterior Production Losses by State (baseline simulation)

The posterior production losses can also be grouped by industry sector, $Q_i = \tilde{\mathbf{q}}_i^T \tilde{\mathbf{x}}_i$, and the five most impacted sectors are shown in Figure 4.5. The agricultural sector suffers the most production losses, which is not surprising given the large amount of grain that is shipped through the UMR. Additionally, agricultural shipments have a lower value per ton than many of the other waterway shipments, so $C_{port}$ tends to be less than $C_{alt}$ in many cases (i.e., companies elect to not pay the added costs of shipping via rail). In turn, production losses from the agriculture industry affect the food processing industry, which is the sixth most impacted sector (see, Supplementary Details). Chemical processing, metals manufacturing, and petroleum (including coal) are also highly affected sectors, as they comprise much of the shipments that pass through the UMR. Lastly, wholesale retail also suffers significant production losses due to delayed shipments and indirect losses from other, interdependent sectors.

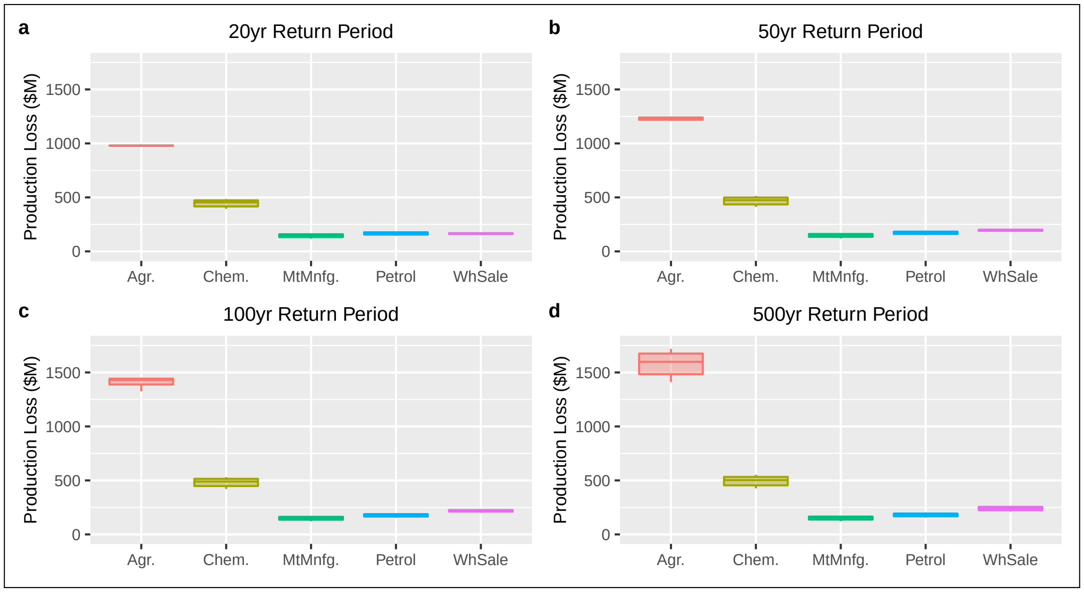



Figure 4.5: Posterior Production Losses by Sector (baseline simulation)

We then simulate production losses from the second version of the model where the Port of Cates Landing is able to handle some of the rerouted shipments, shown in Figure 4.6. As seen, the Port of Cates Landing is able to mitigate regional production losses across all flood scenarios, but its savings exhibit diminishing returns as floods become more severe. This is due to the port's daily throughput reaching max capacity more frequently during the more severe floods.

The cost of developing the port is $53M: $41M to build the port and $12M to connect it to nearby railroads (Byrne, 2020; Davis, 2018). This amount is equivalent to the regional savings in production losses from the port that are realized during a "30-year" flood (see, Supplementary Materials for exact details). We estimate the port's break-even point in terms of a flood's return period, as opposed to years, because we do not place as much trust in the impacts simulated from more-frequent, less-severe floods that would readily be included in estimations of annual savings realized by the port (e.g., the savings from a "2-year" flood the first year, plus the savings from a "1.5-year flood" the second year, etc.). Much of our sentiment stems from the fact that our flood scenarios are not validated for shorter return periods and that a "100-year" event is included in the seven years' worth data from which we develop flood scenarios, likely skewing the left side of the Gumbel distributions toward more extreme outcomes. Additionally, outcomes in our ABM are sensitive at return periods of less than five years (see, Supplementary Details).

Savings in production can be broken down by state and industry sector, Figures 4.7 and 4.8 respectively. As seen, Illinois is the main benefactor of the port. Illinois is the state that has the most outbound shipments

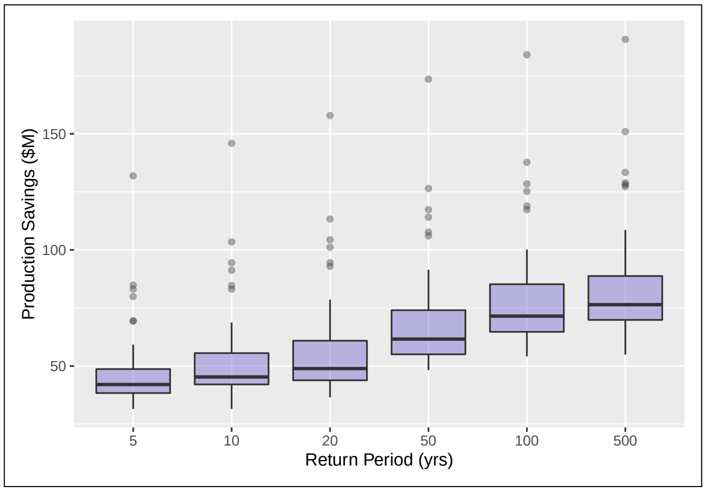



Figure 4.6: Posterior Production Savings from the Port of Cates Landing

through the UMR and is in close proximity to the port, so companies find it cost-effective to reroute their products via rail to the Port of Cates Landing where they can continue via waterway the rest of route. As for industries, chemical manufacturing, petroleum and coal, and agriculture are the main benefactors.

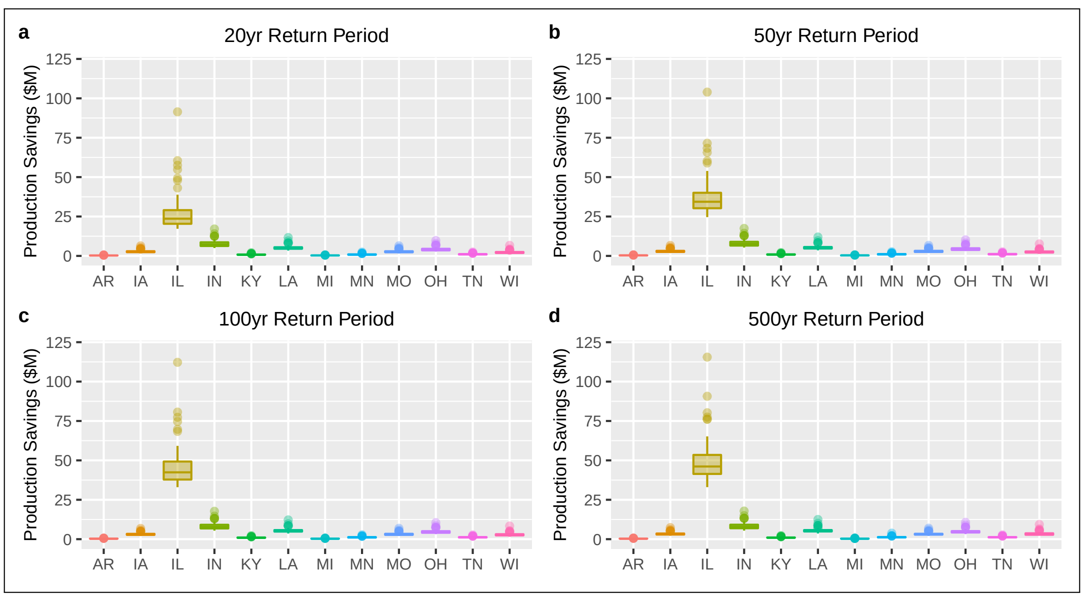



Figure 4.7: Posterior Production Savings by State (Port of Cates Landing)

## 4.4 Discussion

Current risk assessments of disruptions to the inland waterways do not provide guidance on how to empirically specify the likelihood of disruptions occurring, so researchers can use our analysis of an example for how do to so with regards to floods. Using river gauge data in conjunction with WAP procedures serves

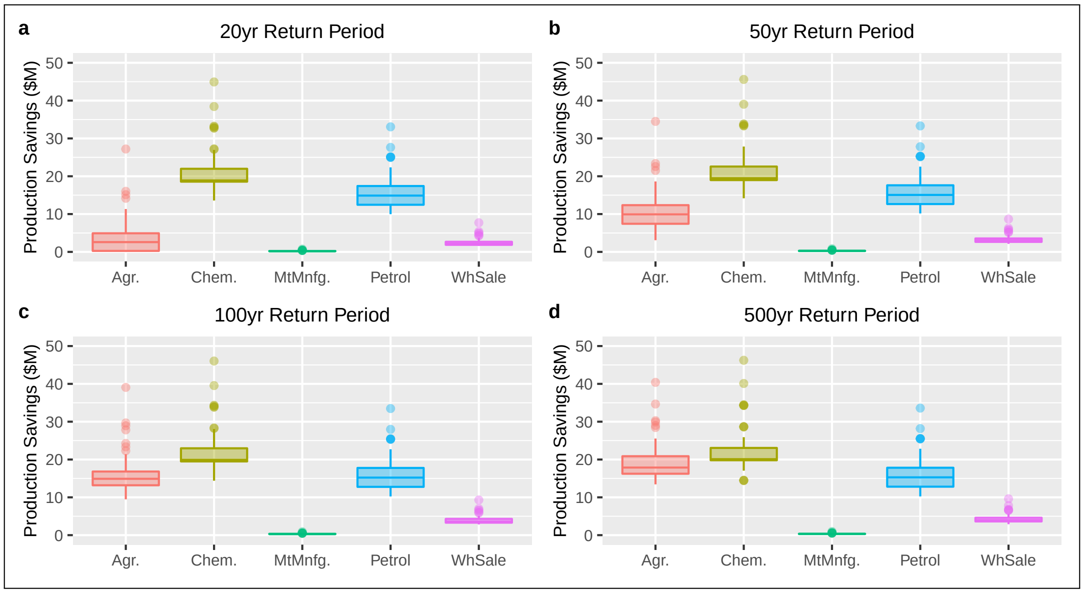



Figure 4.8: Posterior Production Savings by Sector (Port of Cates Landing)

as an intuitive means for modeling the likelihood of river closures, and this approach can be applied to any section of the U.S. inland waterways. Currently, our flood scenarios likely over-estimate the impacts of more-frequent, lower-severity floods because we only have seven years worth of data on which to ground our models, and one of those years includes a "100-year" event. However, the USACE will continue to collect and provide river gauge data, so the accuracy of our forecasts will improve with time. Additionally, it is worth noting that with climate change expected to increase the frequency and severity of extreme weather events, so return periods of floods may be vastly different going forward (e.g. a previously deemed "100-year" flood may now occur every 20 years) (Camp et al., 2013; Pregnolato et al., 2017).

ABC SMC also appears to be well suited for calibrating ABMs involving transportation disruption analyses. These models typically feature complex simulation logic that is difficult to parameterize via traditional methods, and observational data pertaining to impacts from extreme weather events is generally hard to come by (Mahmassani, 1990; He and Liu, 2012; Jackson et al., 2016). The ABC SMC approach is useful under both of these conditions, as we demonstrated here by using the ABC PMC algorithm to calibrate a complex model with only one observed value for reference.

There are several additions to our ABM we would like to include in future iterations. Ideally, we would like to have port-level commodity flow data to make better assumptions about barge availability and to more accurately estimate shipping routes. Similarly, the effects of other economic factors can be explored, including but not limited to port revenue streams, the role of government subsidies, variations in commodity prices due to supply and demand shocks, and contractual agreements between businesses and shipping companies.

However, in its current form, our ABM can help inform the private sector about how their business decisions effect the rest of the economy and give policy makers a better idea of how public waterborne infrastructure (e.g., the Port of Cates Landing) may be a cost effective means of mitigating regional production losses resulting from waterway flood closures.

Regarding this savings in regional production losses, our case study with the Port of Cates Landing highlights a potential cooperative dilemma between the involved states. The production savings from the port are not distributed evenly between states, and depending how the port is funded, these benefits could exist as positive externalities to some states (i.e., states may benefit from the port without contributing to its development). Currently, Tennessee and Kentucky are the only two states that have contributed funding to the port (Byrne, 2020; Arik and Penn, 2009).

### 4.5 Conclusion

The U.S. inland waterways play a vital role in the domestic economy, and extreme weather events, floods in particular, threaten to disrupt operations along the inland waterways. We were able to extend an existing ABM that simulates impacts of disruptions to the inland waterway system by using a data-driven approach to incorporate the likelihood of flood disruptions along the UMR and by calibrating model parameters via ABC SMC (MacKenzie et al., 2012). In turn, we were able to assess corresponding economic risks to the UMR region and explore potential opportunities for mitigation that involve a publicly operated, flood-resilient port (i.e., the Port of Cates Landing) being able to handle rerouted shipments during the disruptions.

We find that Illinois, Louisiana, Minnesota, and Missouri are the states that suffer the most production losses from waterway disruptions along the UMR and that agriculture and chemical manufacturing are the most impacted sectors. Additionally, we find that the cost of developing the port is equivalent to the amount of regional production savings realized from the port during a “30-year” flood event.

### 4.6 Supplementary Materials

The annotated source code for our ABM and all accompanying data can be found on Open Science Framework: https://osf.io/g7xeq/?view_only=51724c6861c44584b34eb1fac5659d01

# CHAPTER 5

## Predicting interstate cooperation in the development of flood-resilient waterborne infrastructure: a common-pool resource dilemma

Community resilience plays a significant role in determining the risks associated with natural disasters, and investments in public infrastructure that bolster community resilience often engender problems of collective coordination typical of cooperative dilemmas. We develop an agent-based model in the form of a common-pool resource dilemma that simulates how decision-makers from states along the Mississippi River may reasonably decide to invest in the development of a publicly operated flood-resilient port that can lead to savings in economic production losses suffered during inland waterway disruptions due to floods. Decisions are informed by an empirically-based decision model trained on the results of a similarly themed behavioral study and suggest that state decision-makers will tend to overvalue the monetary returns from the port. As such, these decisions lead to inefficient and what could be construed as unfair monetary outcomes for many of the involved states, thereby reducing the chances for interstate cooperation in the future. However, when the decision model is viewed as a utility function, outcomes can be presented in terms of subjective values realized by each state, and paths toward resolving the dilemma are revealed.

### 5.1 Introduction

Community resilience to natural disasters is a function of and often strengthened by investments in public infrastructure and related projects (e.g., public hospitals, sea walls, and levees). These projects are generally viewed as common-pool resources (CPRs) or public goods (PGs), depending on the rivalrous nature of their consumption (Ayyub et al., 2016). In either case, problems of collective coordination arise because some individuals may benefit from these developments without investing in them (i.e., free-ride).

There is a rich history of research on social dilemmas and the sustainable management of common resources (Hardin, 1968; Ostrom, 1999; Wunder et al., 2013; Brick et al., 2016). Additionally, there is a growing body of literature on game-theoretic applications to disaster management (Seaberg et al., 2017; Adida et al., 2011; Kunreuther and Michel-Kerjan, 2015; Bouzat and Kuperman, 2014). However, comparatively little work has been done for investigating decision-makers' tendencies to invest in projects that bolster community resilience through the lens of a PG or CPR game.

We construct an agent-based model (ABM) that simulates the dynamics of a one-shot CPR dilemma pertaining to the development of a publicly operated, flood-resilient port along the US inland waterways, namely the Port of Cates Landing. In Chapter 4, we demonstrated that this port has the potential to reduce economic production losses due to floods for states that have significant amounts of inbound/outbound shipments through the Upper Mississippi River (UMR). However, these savings exist as externalities depending how the port is funded. Our ABM simulates how decision-makers from states that benefit from this port may reasonably decide to invest its development and how these decisions affect collective economic outcomes. Decisions are based on two different strategies: 1) one that leads to a Pareto efficient outcome for the region and an individually efficient outcome for each state, and 2) one that stems from an empirically-based behavioral model that predicts deviations from the aforementioned Pareto efficient strategy (i.e., the BART decision model from Chapter 3). Results highlight conditions that would likely engender failed cooperation between states but also reveal potential paths toward resolving the dilemma. This exercise serves as a proof of concept of how researches can use empirically-based decision models to analyze and interpret decision-makers' tendencies and biases in CPR or PG settings and quantify the emergent impacts of their decisions.

### 5.2 Background

The U.S. inland waterway system comprises an extensive collection of ports, channels, and locks and dams, and disruptions to its operations can have devastating impacts on the domestic economy (MacKenzie et al., 2012; Camp et al., 2013; Pregnolato et al., 2017). Despite these known risks, little research has been devoted to analyzing effects of disruptions along the inland waterways (Folga et al., 2009; MacKenzie et al., 2012; Darayi et al., 2019). Recently, MacKenzie et al. (2012) developed an ABM that simulates how companies affected by inland waterway port closures may decide to reroute their shipments and calculates the subsequent economic impacts of these decisions via a multiregional inoperability input-output model (MRIIM) (Crowther and Haimes, 2009; Santos, 2006; Santos and Haimes, 2004; Leontief, 1936).

In Chapter 4, we extended MacKenzie et al. (2012)'s framework to simulate impacts of disruptions due to floods along the UMR. The UMR is a key segment of the U.S. inland waterways, and every year, more than 85 different commodities totalling 119 million tons are transported through this section of the Mississippi River (Fahie, 2019). We were able to simulate regional production losses resulting from various flood events along the UMR and estimate savings in cases where a public port that operates above the 500-year floodplain, the Port of Cates Landing, could have feasibly rerouted some of the disrupted shipments.

However, we also found that the production savings from the port are not distributed evenly between states, and depending how the port is funded, these benefits could exist as positive externalities to some states (see, Figure 4.7). Putting aside potential revenue streams and/or jobs created by the port, this dynamic describes a CPR dilemma. States contribute varying amounts to develop the infrastructure, but the entire region benefits from savings in economic production losses. Repeated over time, interactions such as this one could lead to failed cooperation between states for similar projects in the future (Ostrom, 1999; Schindler, 2012). As such, we simulate how decision-makers from these states that have significant amounts of inbound/outbound shipments through the UMR may hypothetically respond in this particular cooperative dilemma.

There have been other studies that modeled aspects of disaster resilience in a PG or CPR setting. Bednarik et al. (2019) developed a dynamic public goods game (PGG) where forest resources that are exploited for profit also used to protect individuals from flood damages; however, they did not explore the extent to which this added feature causes participants to deviate from behavioral norms in cooperative settings. Additionally, several studies have conceptualized climate change mitigation as a PGG (Hasson et al., 2010; Brick et al., 2016; Kumar and Dutt, 2015); similar to other cooperative dilemmas, results suggest that communication between decision-makers helps promote pro-social behaviors and that the heterogeneity of decision-makers significantly affects outcomes.

Since it was found that individuals in laboratory and real-life settings often do not follow the rational, selfish path toward the destruction of common resources as presumed by Hardin (1968) (i.e., the "Tragedy of the Commons"), much work has been devoted to better understanding how and why individuals behave the way they do in cooperative dilemmas. Individuals' attitudes toward concepts including but not limited to reciprocity, trust, fairness, risk, conformity, and inequity can affect their decisions and collective outcomes (Schindler, 2012; Johnson, 2015; Dong et al., 2016). As such, by measuring differences in individuals' dispositions, it is presumed that researchers can identify characteristics that predict decision-making patterns in these settings (Hirsh and Peterson, 2009). In the field of psychometrics, the Big Five model of personality is the consensus framework for quantifying individuals' differences in temperament and proposes that these differences can be measured along five dimensions (i.e., traits): *Openness*, *Conscientiousness*, *Extraversion*, *Agreeableness*, and *Neuroticism* (John and Srivastava, 1999). Previous research confirms that the Big Five traits are predictive of various response patterns in cooperative dilemmas (Zhao and Smillie, 2015; Kline et al., 2019).

In Chapter 3, using participants' Big Five traits and in-game circumstances as input variables, we constructed Bayesian additive regression tree (BART) decision models to accurately predict deviations from Pareto efficient strategies in a CPR-game that framed player contributions as investments toward public infrastructure that bolstered community resilience. We utilize this BART model to simulate how decision-makers may reasonably respond to different flood scenarios in our ABM. Recently, there has been an increasing drive among researchers to incorporate more realistic, empirically-based decision models in ABMs (Janssen and Ostrom, 2006; Abdulkareem et al., 2018; Choi and Lee, 2018; Nay and Gilligan, 2015). Our study serves as a proof of concept of how researchers can do so for applied CPR or PG dilemmas.

## 5.3 Methods: Agent-based Model

We develop an ABM that simulates how decision-makers from states along the Mississippi River may hypothetically respond in a cooperative dilemma involving the development of a publicly operated, flood-resilient port. Below, we use the ODD + D protocol to describe our model (Müller et al., 2013).

### 5.3.1 Overview

**Purpose**

The purpose of this ABM is to demonstrate a proof of concept of how researchers can use empirically-based decision models to better analyze decisions and resulting outcomes in cooperative dilemmas, in particular those pertaining to aspects of community resilience. We illustrate this notion with a hypothetical CPR dilemma involving the Port of Cates Landing, a public port whose operations are unique in that they lie above the "500-year" floodplain and can conceivably reroute some of the southbound waterway shipments from the UMR during severe flood disruptions (see, Figure 4.1). Our ABM is primarily designed for academic researchers who are interested in game theoretic applications for public infrastructure investments, but policy-makers may also find utility in its insights.

**Entities, state variables and scales**

Our model features one entity type, a state (i.e., the decision-maker representing a state). There are 12 states that have significant amounts of outbound and/or inbound shipments through the UMR, as revealed by USACE's Waterborne Commerce Statistics Center (WCSC) (USACE, 2017). These states are Minnesota, Wisconsin, Iowa, Missouri, Illinois, Indiana, Ohio, Kentucky, Tennessee, Arkansas, Mississippi, and Louisiana (see, Figure 4.1). Each state has the following attributes, as shown in Table 5.1: available funding (estimated as the portion of state tax revenues that are expected to be allocated to transportation related expenditures over a five-year planning period), and average state-wide assessments of Big Five personality inventories (FTA, 2019; Chantrill, 2021; Rentfrow et al., 2008). Given these attributes, decision-makers representing each state decide how much funding they want to allocate to the development of the flood-resilient port in response to various flood scenarios.

There are three exogenous factors in our model. The first is the flood scenario, which consists of the number of days per year that segments of the UMR are expected to be closed due to high-water conditions. The second exogenous factor is the planning period over which the value the flood-resilient port are tabulated. Here, we set the planning period to be five years; this choice is arbitrary and simply used to demonstrate our approach as a proof-of-concept. The last exogenous factor is the cost of developing the flood-resilient port. Here, the estimated cost for developing the Port of Cates Landing is \$53M (Byrne, 2020; Davis, 2018). If

Table 5.1: State Input Variables: Available Funds and Big Five Traits

| | Funds ($M) | O (z-score) | C (z-score) | E (z-score) | A (z-score) | N (z-score) |
|---|---|---|---|---|---|---|
| Arkansas | 2.8 | 0.06 | -0.54 | -0.21 | -0.53 | 1.01 |
| Illinois | 13.7 | 0.17 | 0.90 | 0.80 | 0.07 | 0.21 |
| Indiana | 6.7 | -0.18 | 0.71 | -0.36 | 0.38 | 0.88 |
| Iowa | 3.1 | -0.97 | -0.44 | 0.45 | 0.54 | 0.15 |
| Kentucky | 4.5 | -1.10 | 0.37 | -0.41 | 0.22 | 1.17 |
| Louisiana | 4.0 | -0.01 | -0.24 | -0.20 | 0.55 | 1.14 |
| Minnesota | 8.3 | -0.67 | 0.14 | 1.29 | 1.41 | -0.80 |
| Mississippi | 2.4 | -0.80 | 0.79 | 0.34 | 1.39 | 1.50 |
| Missouri | 2.4 | -0.04 | 0.97 | 0.35 | 0.51 | -0.09 |
| Ohio | 6.7 | 0.12 | -0.56 | -0.05 | 0.04 | 1.10 |
| Tennessee | 5.9 | 0.19 | 0.72 | -0.19 | 1.08 | 0.11 |
| Wisconsin | 6.9 | -1.31 | 0.29 | 2.14 | 1.32 | -0.45 |

users wish to apply our methodology to the development of other ports or infrastructure projects, they would need to change this metric as it influences the states' investment decisions.

**Process overview and scheduling**

Each simulation run consists of three main processes: 1) an initial flood scenario, 2) states' investment decisions in response to this scenario, and 3) the resulting impacts of these decisions simulated over the planning period.

First, a UMR flood scenario is sampled from the empirical flood distributions we developed in Chapter 4 using historical river gauge data and high-water operating procedures outlined in the Waterways Action Plan (WAP) (USACE, 2021; USCG, 2020). The empirical flood distributions predict the number of days per year each segment of the UMR is expected to be closed based on the return period of a flood (e.g., a "100-year" flood, or a flood expected to occur once every 100 years). A scenario is randomly sampled from this distribution and then propagated through the ABM that we also developed in Chapter 4 to determine the subsequent production losses suffered by each state and to estimate the savings that could have been realized had the Port of Cates Landing been operational and able to reroute some of the disrupted shipments.

Given these outcomes, state decision-makers respond to the flood scenario by deciding how much they want to invest in the development of the flood-resilient port. In other words, the flood scenario serves as the impetus for the residents of a state and/or their elected decision-makers to invest in infrastructure that helps protect their communities from future floods. We model decisions in this manner because policy-making tends to be reactionary as opposed to forward-looking. When the impacts from a flood are more severe and the potential savings from the flood-resilient port are greater, residents and/or their representatives are presumed to be more inclined to invest in the infrastructure.

Lastly, impacts resulting from the states' investment decisions are simulated over a predetermined planning period. For each period (i.e., year), a flood scenario is randomly sampled from the empirical flood distribution. If states funded the development of the Port of Cates Landing, outbound shipments from the UMR are able to reroute through it during the disruptions. If not, the full impact of the flood is realized. Costs and benefits are tabulated over the planning period in terms of monetary value and subjective utility for each state, discussed later.

### 5.3.2 Design Concepts

**Theoretical and empirical background**

Our model is predominantly based on two prior studies. First, MacKenzie et al. (2012) developed an ABM that simulates how businesses may reasonably reroute shipments in response to inland waterway port closures and calculates the subsequent economic impacts of such decisions using a multiregional inoperablility input-output model (MRIIM) (Crowther and Haimes, 2009; Santos, 2006; Santos and Haimes, 2004; Leontief, 1936). In the event of a disruption, affected businesses decide whether or not to reroute their product via an alternative shipping method or to leave their products at port, based on a variety of factors which include the duration of the disruption, corresponding shipping costs, and desire of of a business to satisfying customer demand (MacKenzie et al., 2012). If companies ship their product via an alternative method, it is assumed demand is satisfied (MacKenzie et al., 2012). However, if companies decide to keep their product at port, this delay leads to intermediate industries experiencing losses in production (MacKenzie et al., 2012; Horowitz and Planting, 2009).

Second, in Chapter 4, we extended MacKenzie et al. (2012)'s model to incorporate empirically grounded forecasts for the expected number of days sections of the UMR would be closed due to high-water conditions given a flood's return period and high-water operating procedures as documented in the WAP. Additionally, we used approximate Bayesian computation with sequential Monte Carlo sampling (ABC SMC) to calibrate model parameters, which allowed us to simulate valid estimates for regional production losses resulting from various flood scenarios and the savings that could have been realized if the Port of Cates Landing had been available to reroute some of the affected shipments.

**Individual decision-making**

Decision-making is modeled at the state level and is based on one of two different strategies (i.e., two different versions of the simulation). Within each version, all state decision-makers follow the same strategy.

The first strategy leads to a simultaneously Pareto efficient outcome for the region and an individually efficient outcome for each state (i.e., states invest an amount that is proportional to the amount of savings in production losses they would have realized from the Port of Cates Landing being able to reroute shipments during the flood disruption). However, states only invest if the total region's savings meets or exceeds the development costs of the port, which in our case is $53 million. We use this threshold to determine whether there is enough public sentiment in response to flood scenarios to warrant such an expenditure. If there is not enough interest, no investment occurs, and available funds are assumed to be directed elsewhere.

For the second strategy, states invest according to a Bayesian Additive Regression Tree (BART) decision model. This model was trained on the results of a web-based CPR game that framed participant contributions as investments toward the development of infrastructure that bolstered community resilience (see, Chapter 3). The BART model predicts deviations from various Pareto efficient strategies, including the aforementioned one, given the following input variables: the economic circumstances of an individual, the economic circumstances of his or her community, the expected impact of a disaster, the expected benefits that the individual and his or her community realize from the proposed infrastructure, and the individual's temperament as measured by the Big Five personality traits.

We use a corresponding set of variables as inputs to our ABM, as previously shown in Table 5.1. We assume that tax revenues characterize the economic circumstances of the states (FTA, 2019; Chantrill, 2021). We also assume that the state-wide Big Five inventory surveys conducted by Rentfrow et al. (2008) reflect the average disposition of individuals residing in these states, which in turn reflects the disposition of the states' decision-makers. The expected impacts of the disaster (i.e., flood disruption) and expected benefits from the proposed infrastructure (i.e., flood-resilient port) are simulated from the ABM discussed in Chapter 4.

Like with the Pareto efficient strategy, states only invest via this second strategy if the collective interest as predicted by the BART model meets or exceeds the cost of developing the port. In this manner, predictions from the BART model can be conceptualized as each state's perceived utility of the investment (Rustichini et al., 2016).

**Learning**

No individual or collective learning is included in the decision process.

**Individual sensing**

State decision-makers are assumed to know how much funding is available to each state and how much each state benefits from the flood-resilient port. Decision-makers are also aware of all exogenous factors: the initial flood scenario, the cost of developing the port, and the length of time over which benefits of the port will be simulated.

**Individual prediction**

No individual or collective prediction occurs in this simulation.

**Interaction**

Interactions between states are indirect through their investment decisions. If there is enough collective investment in the port, it will be developed and affect all states via savings in production losses.

**Collectives**

States do not belong to or form any collectives.

**Heterogeneity**

States are heterogeneous with respect to their attributes (see, Table 5.1) but homogeneous with respect to their decision-making strategies.

**Stochasticity**

Stochasticity is introduced in the generation of flood scenarios. Percentiles pertaining to the return period of the flood scenarios are randomly generated from a uniform distribution with a range of zero to one. Because we used Gumbel distributions to model the flood scenarios in Chapter 4, percentiles can be translated to return periods via the following equation: $return_{period} = 1/(1 - return_{percentile})$ (Gilleland and Katz, 2016).

**Observation**

After each simulation run, we record the amount that states decided to invest in the port using each strategy. We also record the emergent savings in production losses realized over the planning period, both in terms of monetary value and subjective utility.

### 5.3.3 Details

**Implementation details**

Our ABM is implemented in *R* programming language. The source code and accompanying data files can be found in the Supplementary Materials.

**Initialization**

Our model is initialized with a planning period of five years and a total number of 1000 simulation runs.

**Input data**

State input attributes are presented in Table 5.1. The Supplementary Materials contain all other data necessary to recreate our analysis.

**Submodels**

Our simulation includes three submodels: 1) flood scenarios, 2) state investment decisions in response to (1), and 3) simulated economic outcomes of (2) over the planning period. We continuously tested these submodels for congruence between model logic and outcomes as we developed the ABM.

## 5.4 Results

When responding to the majority of flood scenarios, states do not to invest in the port. According to the Pareto efficient strategy, states develop the port in 45 of the 1000 simulation runs (4.5%). According to the BART decision model, which better reflects how individuals behave in a similarly framed CPR game, states develop the port in 304 of the 1000 runs (30.4%). Below, we discuss the results of these decisions in terms of three different outcomes: 1) monetary valuations of the port based on the Pareto efficient strategy, 2) monetary valuations based on how states act according the BART decision model, and 3) valuations measured in terms of subjective utility, based on the same BART model.

Table 5.2 shows the expected monetary valuations of the port when states act according to the Pareto efficient strategy. Cash flows (Y1:Y5) are the average annual savings in production losses across the 1000 simulation runs. The cost of developing the port (C0) reflects the average amount each state contributes to the port in the 45 runs that resulted in the port being developed based on the Pareto efficient strategy. As this Pareto efficient strategy optimizes both the individual and collective return on investment, it is not surprising that return on investment (ROI) and internal rate of return (IRR) for each state and the overall project are all positive. Additionally, these returns are distributed fairly evenly between the states on a percentage basis. Both of these conditions are favorable for sustained interstate cooperation (Dong et al., 2016; Johnson, 2015; Ostrom, 1999).

Table 5.2: Monetary Valuation of Pareto Efficient Strategy ($M)

| | C0 | Y1 | Y2 | Y3 | Y4 | Y5 | ROI | IRR |
|---|---|---|---|---|---|---|---|---|
| Arkansas | (0.24) | 0.07 | 0.08 | 0.08 | 0.08 | 0.07 | 55% | 17% |
| Illinois | (29.7) | 7.37 | 8.02 | 7.68 | 7.51 | 6.81 | 26% | 8% |
| Indiana | (6.28) | 2.41 | 2.61 | 2.51 | 2.48 | 2.23 | 95% | 28% |
| Iowa | (2.29) | 0.70 | 0.76 | 0.73 | 0.72 | 0.66 | 56% | 17% |
| Kentucky | (0.70) | 0.21 | 0.23 | 0.22 | 0.22 | 0.20 | 55% | 17% |
| Louisiana | (4.06) | 1.37 | 1.49 | 1.43 | 1.41 | 1.28 | 72% | 22% |
| Minnesota | (0.85) | 0.23 | 0.25 | 0.24 | 0.23 | 0.21 | 37% | 12% |
| Mississippi | (0.24) | 0.08 | 0.08 | 0.08 | 0.08 | 0.07 | 62% | 19% |
| Missouri | (2.27) | 0.70 | 0.76 | 0.73 | 0.72 | 0.65 | 56% | 17% |
| Ohio | (3.42) | 1.09 | 1.19 | 1.14 | 1.13 | 1.02 | 63% | 19% |
| Tennessee | (0.88) | 0.27 | 0.29 | 0.28 | 0.28 | 0.25 | 55% | 17% |
| Wisconsin | (2.02) | 0.56 | 0.61 | 0.59 | 0.58 | 0.53 | 42% | 13% |
| | (53.0) | 15.06 | 16.37 | 15.70 | 15.40 | 13.97 | 44% | 14% |

Table 5.3 displays monetary valuations of the port given that the states act according to the BART decision

model. Y1:Y5 consist of the same expected savings in production losses that are in Table 5.2. However, C0 reflects the average amount each state contributes in the 304 runs that resulted in the port being developed based on the BART decision model. As seen, most of the states overestimate monetary returns from the port (i.e., they contribute more than what the did in the Pareto efficient strategy). These decisions lead to inefficient and uneven outcomes, with several of the states realizing negative returns. Neither of these conditions is conducive for sustained cooperation (Dong et al., 2016; Johnson, 2015; Ostrom, 1999). Furthermore, these outcomes are more likely to occur than the Pareto efficient outcomes for two reasons: 1) conceptually, the BART decision model better reflects how individuals act in a similarly framed CPR game and 2) empirically, more simulation runs resulted in the port being developed using this strategy than the Pareto efficient one (304 vs. 45). As such, we can likely expect sustained cooperation to fail in this dilemma.

Table 5.3: Monetary Valuation of BART Decision Model ($M)

| | C0 | Y1 | Y2 | Y3 | Y4 | Y5 | ROI | IRR |
|---|---|---|---|---|---|---|---|---|
| Arkansas | (1.06) | 0.07 | 0.08 | 0.08 | 0.08 | 0.07 | -65% | -27% |
| Illinois | (17.3) | 7.37 | 8.02 | 7.68 | 7.51 | 6.81 | 116% | 33% |
| Indiana | (5.67) | 2.41 | 2.61 | 2.51 | 2.48 | 2.23 | 116% | 33% |
| Iowa | (4.32) | 0.70 | 0.76 | 0.73 | 0.72 | 0.66 | -17% | -6% |
| Kentucky | (2.23) | 0.21 | 0.23 | 0.22 | 0.22 | 0.20 | -51% | -20% |
| Louisiana | (4.10) | 1.37 | 1.49 | 1.43 | 1.41 | 1.28 | 70% | 20% |
| Minnesota | (2.38) | 0.23 | 0.25 | 0.24 | 0.23 | 0.21 | -51% | -20% |
| Mississippi | (0.65) | 0.08 | 0.08 | 0.08 | 0.08 | 0.07 | -39% | -15% |
| Missouri | (2.41) | 0.70 | 0.76 | 0.73 | 0.72 | 0.65 | 47% | 15% |
| Ohio | (5.47) | 1.09 | 1.19 | 1.14 | 1.13 | 1.02 | 2% | 0% |
| Tennessee | (3.08) | 0.27 | 0.29 | 0.28 | 0.28 | 0.25 | -56% | -22% |
| Wisconsin | (4.32) | 0.56 | 0.61 | 0.59 | 0.58 | 0.53 | -34% | -12% |
| | (53.0) | 15.06 | 16.37 | 15.70 | 15.40 | 13.97 | 44% | 14% |

However, failures of cooperative dilemmas can often be viewed as individuals failing to communicate different utility functions to one another (Ostrom, 1999; Johnson, 2015). As such, when valuations of the port are measured in terms of subjective utility defined by the same BART decision model, potential paths toward reconciling this dilemma are revealed. Table 5.4 contains the same values for C0 that are in Table 5.3. However, Y1:Y5 are expressed as the average annual investments that states would have hypothetically made in response to the flood scenarios simulated across all runs based on the strategy informed by the BART model. In other words, Y1:Y5 express the average subjective value that each state derives from the port each year. When valuations are viewed in this manner, the perceived return on investments for most states is again positive, with only Mississippi and Missouri deriving negative utility from the investment.

By considering the subjective utility of the investment, states can better understand the motivation behind one another's decision-making tendencies. In turn, this understanding can help facilitate dialogue between state decision-makers and improve the chances for cooperation (Dong et al., 2016; Johnson, 2015). For example, most states ostensibly overvalue the monetary returns from the port (see, Table 5.3), but they may

Table 5.4: Subjective Valuation of BART Decision Model ($M)

| | C0 | Y1 | Y2 | Y3 | Y4 | Y5 | ROI | IRR |
|---|---|---|---|---|---|---|---|---|
| Arkansas | (1.06) | 0.45 | 0.25 | 0.26 | 0.20 | 0.19 | 27% | 10% |
| Illinois | (17.3) | 7.33 | 7.84 | 7.58 | 7.42 | 6.87 | 114% | 33% |
| Indiana | (5.67) | 2.39 | 2.31 | 2.25 | 2.19 | 2.01 | 96% | 29% |
| Iowa | (4.32) | 1.29 | 1.15 | 1.14 | 1.09 | 1.04 | 32% | 10% |
| Kentucky | (2.23) | 1.50 | 1.33 | 1.33 | 1.27 | 1.26 | 200% | 55% |
| Louisiana | (4.10) | 1.88 | 1.73 | 1.71 | 1.65 | 1.57 | 109% | 32% |
| Minnesota | (2.38) | 1.47 | 1.30 | 1.39 | 1.33 | 1.32 | 186% | 51% |
| Mississippi | (0.65) | (0.11) | (0.24) | (0.24) | (0.30) | (0.30) | -283% | - |
| Missouri | (2.41) | (0.01) | (0.17) | (0.18) | (0.23) | (0.27) | -136% | - |
| Ohio | (5.47) | 3.29 | 3.14 | 3.12 | 3.06 | 2.98 | 185% | 51% |
| Tennessee | (3.08) | 1.56 | 1.38 | 1.38 | 1.33 | 1.31 | 126% | 37% |
| Wisconsin | (4.32) | 2.72 | 2.59 | 2.67 | 2.62 | 2.58 | 205% | 55% |
| | (53.0) | 23.77 | 22.60 | 22.39 | 21.64 | 20.57 | 109% | 32% |

actually just be deriving greater utility from the guaranteed savings of the port over risking impacts without it (see, Table 5.4). In other words, states are risk-averse. Similarly, the negative utility experienced by Mississippi and Missouri stems from the fact that after controlling for differences in temperament, they are among the states with the least amount of funding available to contribute. With this knowledge and due to the fact that many individuals are also averse to inequities, other states may be more inclined to absolve Mississippi's and Missouri's negative sentiments toward the project and/or any decisions to free-ride in the future (Zhao and Smillie, 2015).

It is often difficult to distinguish subjective utility from errors in human judgment when characterizing deviations from choices that do not satisfy rational value maximization (Amador-Hidalgo et al., 2021). However, even if predictions from the BART model are interpreted more as errors in human judgment, as opposed to subjective utility, they can still facilitate communication by informing state decision-makers of their potential biases. For example, as mentioned, many states overvalue the monetary returns from the port and consequently realize negative outcomes (see, Table 5.3). If state decision-makers are aware of their propensity to over-contribute, they can adjust their actions accordingly if needed and arrive at a more effective solution (e.g., the Pareto efficient outcome in Table 5.2).

## 5.5 Discussion

Our ABM demonstrates a proof of concept for how researchers can use empirically grounded decision models to better understand and interpret decisions in an applied CPR or PG setting. Predictions from the decision models can be interpreted as subjective valuations or as errors in human judgment. In either case, these models reveal the tendencies and biases of decision-makers and also help quantify how decisions lead to various outcomes. As such, this type of exercise can help facilitate dialogue between decision-makers, which in turn improves the chances for sustained cooperation in the management of common resources (Dong et al., 2016; Johnson, 2015).

This type of approach assumes that decision models based on the results of laboratory studies can effectively be applied to real-life settings. For example, in Chapter 3, we were able to develop a decision model that accurately predicted individuals' responses in a similarly-themed CPR game. However, there is no guarantee that this model can accurately predict realistic investment decisions in our applied ABM. In particular, we assume that differences in individuals' dispositions as measured by Big Five traits can be applied to decision-making at the state level. This assumption may be reasonable in a democratic republic like the U.S. where individuals elect state officials and given the fact that politicians have analyzed their constituents' Big Five traits to tailor campaign messages (Kosinski et al., 2013; Isaak and Hanna, 2018). However, it is nonetheless an assumption and potential limitation of our analysis.

Lastly, our work further demonstrates the potential benefits of incorporating aspects of psychometrics with formal decision theory, and in recent years, scientists have recognized the intuitive overlap between these fields (Bortoli et al., 2019; Boyce et al., 2016; Rustichini et al., 2016; Schröder and Gilboa Freedman, 2020). For example, Boyce et al. (2019) found success in using Big Five inventories to predict public sentiment among Estonian and Latvian residents toward investing in various public environmental goods. In this manner, similar types of inquiries could be conducted in CPR or PG settings to help construct multi-utility decision models that better interpret stakeholders' valuations of projects in terms of monetary, environmental, and social impacts.

### 5.6 Conclusion

Community resilience plays a significant role in determining the risks associated with natural disasters, and investments in public infrastructure that increase community resilience often engender cooperative dilemmas associated with CPRs or PGs. We develop an ABM in the form of a one-shot CPR dilemma that simulates how states along the Mississippi River may decide to invest in a flood-resilient port that can lead to savings in production losses suffered during inland waterway disruptions due to floods. We compare these decisions and resulting outcomes based on the states operating according to two different strategies: 1) one that leads to a simultaneous Pareto efficient outcome for the region and an individually efficient outcome for each state, and 2) one informed by an empirically grounded decision model trained on the results of a similarly themed behavioral study.

When operating according to the Pareto efficient strategy, states only invest in the port in 45 of the 1000 simulation rums, but the expected returns on investment are positive for each state and distributed fairly evenly on a percentage basis. When states act as predicted by the empirically-based decision model, they invest in the port more frequently (304 of the 1000 runs), but their decisions often lead to inefficient monetary outcomes that are not conducive to sustained cooperation. However, when these outcomes are instead viewed as the states' subjective valuations of the port, as determined by the same decision model, state decision-makers can better understand the motivation behind one another's decision-making tendencies and biases, and paths toward resolving the dilemma are revealed.

### 5.7 Supplementary Materials

The annotated source code for our ABM and all accompanying data can be found on Open Science Framework: https://osf.io/tzckg/?view_only=75cedd91d4aa49929544721dd5897e19

# CHAPTER 6

## Conclusion

As mentioned throughout this dissertation, resilience is an important concept in the context of natural disasters because it helps reconcile why similar events can have such disparate impacts on various communities (Fjord and Manderson, 2009; Johnson et al., 2020). Although research on natural hazards has advanced substantially over the last several decades, comparatively less progress has been made on quantifying community resilience and behavioral uncertainty in the context of these events (UN/ISDR, 2005). In particular, there have been a proliferation of unproven frameworks that purport to measure community resilience without an objective means of comparing and evaluating them (Bakkensen et al., 2017; Johnson et al., 2020). Similarly, agent-based models that attempt to simulate impacts of natural disasters on communities rarely have practical use because they fail to incorporate or justify how human decision-making affects community resilience (Fiedrich and Burghardt, 2007; Janssen and Ostrom, 2006; Taylor et al., 2014). Here, we have made significant contributions to advance the knowledge in both of these areas.

In Chapter 2, we used factor analysis to establish an intuitive and data-driven schema for comparing a popular class of frameworks aimed at measuring community resilience, called resilience indices. Our findings suggest that 50 of the 130 variables that comprise six of the most established indices in the field effectively load onto five main dimensions of community resilience: wealth, poverty, agencies per capita, elderly populations, and non-English speaking populations. This factor model greatly reduces the ambiguity that arises when comparing existing indices and provides a robust framework for evaluating these factors and other variables with respect to their abilities to predict disaster outcomes. As such, these efforts pave the way for researchers to establish construct validity in this field.

In Chapter 3, we developed empirically-based decision models that can be used to predict or generate responses in CPR settings pertaining to public infrastructure investments that bolster community resilience to disasters. Public infrastructure that bolsters community resilience can often be conceptualized as a PG or CPR, so it is important to understand individuals' decision-making patterns in these settings in order to quantify how they can affect collective outcomes. As such, we conducted a behavioral study in the form of a web-based CPR game to elicit individuals' tendencies to contribute to the development of infrastructure that bolsters their communities' resilience to various disaster scenarios. We then trained BART models and MLRs on results of this study to predict deviations from participants' decisions that would lead to Pareto efficient outcomes for their communities and to examine how the Big Five dimensions of personality can predict these decision-making patterns. We find that individuals tend to over-contribute compared to Pareto efficient strategies, indicating a potential link to the concept of risk-aversion from decision theory. However, the BART models are able to accurately predict these deviations and perform much better than their MLR counterparts. In particular, the Big Five trait Openness is a key predictor, and PDPs derived from the BART models reveal that the relationship between Openness and participants' decisions is nonlinear. This finding suggests that it might be prudent to utilize more advanced statistical learning approaches to reexamine assumptions of linearity in other studies that have related individuals' differences in temperaments to responses in cooperative dilemmas.

In Chapter 4, we developed an empirically-based ABM that simulates how businesses along the Mississippi River may reasonably reroute inland waterway shipments in response to various flood scenarios and calculates the subsequent economic impacts of such decisions. In doing so, we extended an existing ABM that analyzes impacts of disruptions to the inland waterway system by empirically modeling the likelihood of flood disruptions and by calibrating model parameters via ABC SMC. To the best of our knowledge, no prior systems-level risk assessment of inland waterway disruptions included empirically grounded assumptions for the likelihood of disruptions occurring or provided guidance on how to empirically calibrate associated models. By addressing these gaps, we were able to assess economic risks to the UMR supply chain and explore potential opportunities for mitigation that involves a publicly operated, flood-resilient port (i.e., the Port of Cates Landing) being able to handle rerouted shipments during flood disruptions. We find that Illinois, Louisiana, Minnesota, and Missouri are the states that suffer the most production losses from waterway closures along the UMR, and agriculture and chemical manufacturing are the most impacted industry sectors. Although the Port of Cates Landing is able to provide relief in the form of savings in production losses, these savings are not distributed evenly between states and exist as externalities depending how the port is funded, thus engendering coordination problems typical of cooperative dilemmas.

In Chapter 5, we combined insights from Chapters 3 and 4 to simulate how state decision-makers may hypothetically contribute funds to the development of the Port of Cates Landing in the aforementioned cooperative dilemma. Based on the BART decision model developed in Chapter 3, state investments often lead to inefficient monetary outcomes that are not favorable for sustained cooperation. However, when decisions are viewed as subjective valuations of the port, decision-makers can better understand the motivation behind one another's tendencies and biases, and paths toward resolving the dilemma are revealed. This exercise serves as a proof of concept of how researchers can use empirically-based decision models to analyze and interpret decision-making patterns in applied CPR or PG settings and quantify the emergent impacts of these decisions.

The work presented in this dissertation reveals several avenues for future research. First, we plan to use statistical learning to relate the factor model established in Chapter 2 to various disaster outcomes. In doing so, we can begin to holistically establish construct validity for resilience indices. Additionally, we plan to systematically vary and study the effects of game-level treatment factors in our web-based CPR study. In turn, we can develop a better empirical understanding of how individuals respond in cooperative dilemmas framed in terms of collective efforts to bolster community resilience. Related, we also plan to use advanced machine learning procedures to reexamine findings from other studies that have assumed linear relationship between individuals' differences in temperaments and responses in cooperative dilemmas. Lastly, we aim to use insights drawn from these aforementioned analyses to develop data-driven, multi-attribute decision models that can be used to better evaluate returns from infrastructure developments in terms of economic, social, and environmental benefits to communities.

## References


Abdulkareem, S. A., Augustijn, E.-W., Mustafa, Y. T., and Filatova, T. (2018). Intelligent judgements over health risks in a spatial agent-based model. *International Journal of Health Geographics*, 17(1):8.

Adger, W. N. (2006). Vulnerability. *Global Environmental Change*, 16(3):268–281.

Adida, E., DeLaurentis, P.-C. C., and Lawley, M. A. (2011). Hospital stockpiling for disaster planning. *IIE Transactions*, 43(5):348–362.

Ahlqvist, O., Khodke, N., and Ramnath, R. (2018). GeoGame analytics – A cyber-enabled petri dish for geographic modeling and simulation. *Computers, Environment and Urban Systems*, 67:1–8.

Alexander, D. (2002). The Study of Natural Disasters, 1977–97: Some Reflections on a Changing Field of Knowledge. *Disasters*, 21(4):284–304.

Allport, G. W. and Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, 47(1):i–171.

Amador-Hidalgo, L., Brañas-Garza, P., Espín, A. M., García-Muñoz, T., and Hernández-Román, A. (2021). Cognitive abilities and risk-taking: Errors, not preferences. *European Economic Review*, 134:103694.

Anderson, K. D. (2010). *Framing Traits: The Role of Personality in Framing Effects*. PhD thesis, University of Nebraska, Lincoln, Nebraska.

Anjoran, R. (2009). Using penalties effectively. Section: Supplier Management.

AOPOA (2010). Shipping on the Waterway - Frequently Asked Questions.

Arik, M. and Penn, D. (2009). Contributions of Proposed Investment in the Northwest Tennessee Regional Port at Cates Landing to the Regional Economy. Technical report, Middle Tennessee State University, Business and Economic Research Center.

Associated Press (2019). When '100-year floods' happen often, what should you call them?

Aven, T. (2016). Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research*, 253(1):1–13.

Axelrod, R. and Hamilton, W. (1981). The evolution of cooperation. *Science*, 211(4489):1390.

Ayyub, B. M. (2015). Practical Resilience Metrics for Planning, Design, and Decision Making. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 1(3):04015008.

Ayyub, B. M., Chapman, R. E., Galloway, G. E., and Wright, R. N. (2016). Economics of Community Disaster Resilience Workshop Proceedings. Technical Report NIST SP 1600, National Institute of Standards and Technology.

Bakkensen, L. A., Fox-Lent, C., Read, L. K., and Linkov, I. (2017). Validating Resilience and Vulnerability Indices in the Context of Natural Disasters. *Risk Analysis*, 37(5):982–1004.

Beaumont, M. A. (2019). Approximate Bayesian Computation. *Annual Review of Statistics and Its Application*, 6(1):379–403.

Beaumont, M. A., Cornuet, J.-M., Marin, J.-M., and Robert, C. P. (2009). Adaptive approximate Bayesian computation. *Biometrika*, 96(4):983–990. arXiv: 0805.2256.

Bednarik, P., Linnerooth-Bayer, J., Magnuszewski, P., and Dieckmann, U. (2019). A Game of Common-pool Resource Management: Effects of Communication, Risky Environment and Worldviews. *Ecological Economics*, 156:287–292.

Ben-Ner, A. and Kramer, A. (2011). Personality and altruism in the dictator game: Relationship to giving to kin, collaborators, competitors, and neutrals. *Personality and Individual Differences*, 51(3):216–221. Place: Netherlands Publisher: Elsevier Science.

Bhagat, N. (2017). Flood Frequency Analysis Using Gumbel's Distribution Method: A Case Study of Lower Mahi Basin, India. *Journal of Water Resources and Ocean Science*, 6(4):51.

Birkmann, J. (2007). *Measuring Vulnerability to Natural Hazards: towards disaster resilient societies*. The Energy and Resources Institute (TERI), New Delhi, India.

Birkmann, J., Kienberger, S., and Alexander, D. E. (2014). Introduction Vulnerability: a key determinant of risk and its importance for risk management and sustainability. In *Assessment of Vulnerability to Natural Hazards*, pages ix–xiii. Elsevier.

Bortoli, D. D., da Costa Jr, N., Goulart, M., and Campara, J. (2019). Personality traits and investor profile analysis: A behavioral finance study. *PLOS ONE*, 14(3):e0214062. Publisher: Public Library of Science.

Bouzat, S. and Kuperman, M. N. (2014). Game theory in models of pedestrian room evacuation. *Physical Review E*, 89(3):032806.

Boyce, C., Czajkowski, M., and Hanley, N. (2019). Personality and economic choices. *Journal of Environmental Economics and Management*, 94:82–100.

Boyce, C. J., Wood, A. M., and Ferguson, E. (2016). Individual Differences in Loss Aversion. *Personality and Social Psychology Bulletin*, page 14.

Brick, K., Visser, M., and Van der Hoven, Z. (2016). Cooperation and Climate Change: Can Communication Facilitate the Provision of Public Goods in Heterogeneous Settings? *Environmental and Resource Economics*, 64(3):421–443.

Brocklebank, S., Lewis, G. J., and Bates, T. C. (2011). Personality accounts for stable preferences and expectations across a range of simple games. *Personality and Individual Differences*, 51(8):881–886. Place: Netherlands Publisher: Elsevier Science.

Burton, C. G. (2015). A Validation of Metrics for Community Resilience to Natural Hazards and Disasters Using the Recovery from Hurricane Katrina as a Case Study. *Annals of the Association of American Geographers*, 105(1):67–86.

Byrne, S. (2020). BUILD Grant Funding Includes Awards Along Inland Waterways. Section: News.

Camp, J., Abkowitz, M., Hornberger, G., Benneyworth, L., and Banks, J. C. (2013). Climate Change and Freight-Transportation Infrastructure: Current Challenges for Adaptation. *Journal of Infrastructure Systems*, 19(4):363–370.

Capraro, V. (2013). A Model of Human Cooperation in Social Dilemmas. *PLoS ONE*, 8(8):e72427.

Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, 1(2):245–276.

Center for Emergency Management and Homeland Security (2018). The Spatial Hazard Events and Losses Database for the United States, Version 17.0 [Online Database]. Technical report, Arizona State University, Phoenix, AZ.

Chan, E. Y. Y. (2013). Bottom-up disaster resilience. *Nature Geoscience*, 6(5):327–328.

Chantrill, C. (2021). US Government Finances: Spending, Deficit, Debt, Revenue since 1792.

Chen, D. L., Schonger, M., and Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

Chipman, H. A., George, E. I., and McCulloch, R. E. (2010). BART: Bayesian additive regression trees. *The Annals of Applied Statistics*, 4(1):266–298.

Choi, B. and Lee, S. (2018). An Empirically Based Agent-Based Model of the Sociocognitive Process of Construction Workers' Safety Behavior. *Journal of Construction Engineering and Management*, 144(2):04017102.

Cimellaro, G. P., Renschler, C., Reinhorn, A. M., and Arendt, L. (2016). PEOPLES: A Framework for Evaluating Resilience. *Journal of Structural Engineering*, 142(10):04016063.

Cisewski, J. (2014). Approximate Bayesian Computing using Sequential Sampling.

Cohen, O., Leykin, D., Lahad, M., Goldberg, A., and Aharonson-Daniel, L. (2013). The conjoint community resiliency assessment measure as a baseline for profiling and predicting community resilience for emergencies. *Technological Forecasting and Social Change*, 80(9):1732–1741.

Cox, R. S. and Hamlen, M. (2015). Community Disaster Resilience and the Rural Resilience Index. *American Behavioral Scientist*, 59(2):220–237.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3):297–334.

Crowther, K. G. and Haimes, Y. Y. (2009). Development of the multiregional inoperability input-output model (MRIIM) for spatial explicitness in preparedness of interdependent regions. *Systems Engineering*, pages 1–19.

Crump, M. J. C., McDonnell, J. V., and Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research. *PLoS ONE*, 8(3):e57410.

Cutter, S. L. (2016). The landscape of disaster resilience indicators in the USA. *Natural Hazards*, 80(2):741–758.

Cutter, S. L., Ash, K. D., and Emrich, C. T. (2014). The geographies of community disaster resilience. *Global Environmental Change*, 29:65–77.

Cutter, S. L., Barnes, L., Berry, M., Burton, C., Evans, E., Tate, E., and Webb, J. (2008). A place-based model for understanding community resilience to natural disasters. *Global Environmental Change*, 18(4):598–606.

Cutter, S. L., Boruff, B. J., and Shirley, W. L. (2003). Social Vulnerability to Environmental Hazards. *Social Science Quarterly*, 84(2):242–261.

Cutter, S. L., Burton, C. G., and Emrich, C. T. (2010). Disaster Resilience Indicators for Benchmarking Baseline Conditions. *Journal of Homeland Security and Emergency Management*, 7(1).

Darayi, M., Barker, K., and Nicholson, C. D. (2019). A multi-industry economic impact perspective on adaptive capacity planning in a freight transportation network. *International Journal of Production Economics*, 208:356–368.

Davis, R. (2018). The Port of Cates Landing: Yesterday, Today and Tomorrow. Technical report, Davis Wealth Services.

DeYoung, C. G., Quilty, L. C., and Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, 93(5):880–896.

Diekert, F. K. (2012). The Tragedy of the Commons from a Game-Theoretic Perspective. *Sustainability*, 4(8):1776–1786.

Dong, Y., Zhang, B., and Tao, Y. (2016). The dynamics of human behavior in the public goods game with institutional incentives. *Scientific Reports*, 6(1):28809.

Ebisudani, M. and Tokai, A. (2017). The Application of Composite Indicators to Disaster Resilience: A Case Study in Osaka Prefecture, Japan. *Journal of Sustainable Development*, 10:81.

Fahie, M. (2019). Impacts of the 2019 Upper Mississippi River Flooding on Barge Movements in the Upper Midwest Region.

Fasiolo, M., Pya, N., and Wood, S. N. (2016). A Comparison of Inferential Methods for Highly Nonlinear State Space Models in Ecology and Epidemiology. *Statistical Science*, 31(1):96–118.

Fekete, A. (2019). Social Vulnerability (Re-)Assessment in Context to Natural Hazards: Review of the Usefulness of the Spatial Indicator Approach and Investigations of Validation Demands. *International Journal of Disaster Risk Science*, 10(2):220–232.

Fiedrich, F. and Burghardt, P. (2007). Agent-based systems for disaster management. *Communications of the ACM*, 50(3):41.

Fjord, L. and Manderson, L. (2009). Anthropological Perspectives on Disasters and Disability: An Introduction. *Human Organization*, 68(1):64–72.

Flanagan, B. E., Gregory, E. W., Hallisey, E. J., Heitgerd, J. L., and Lewis, B. (2011). A Social Vulnerability Index for Disaster Management. *Journal of Homeland Security and Emergency Management*, 8(1).

Folga, S., Allison, T., Seda-Sanabria, Y., Matheu, E., Milam, T., Ryan, R., and Peerenboom, J. (2009). A systems-level methodology for the analysis of inland waterway infrastructure disruptions. *Journal of Transportation Security*, 2(4):121–136.

Folger, J. (2018). The Most Popular States to Retire to in the U.S.

Foster, K. (2012). In Search of Regional Resilience. In Weir, M., Pindus, N., Wial, H., and Wolman, H., editors, *Urban and Regional Policy and Its Effects: Building Resilient Regions*, pages 24–59. Brookings Institution Press, Washington DC, UNITED STATES.

Frankenberger, T., Mueller, M., Spangler, T., and Alexander, S. (2013). Community Resilience: Conceptual Framework and Measurement Feed the Future Learning Agenda. Technical report, USAID.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5):1189–1232.

FTA (2019). 2019 State Tax Revenue.

Gasparini, N. and Yuill, B. (2020). High Water: Prolonged Flooding on the Deltaic Mississippi River.

Gigerenzer, G. (1996). The Psychology of Good Judgment: Frequency Formats and Simple Algorithms. *Medical Decision Making*, 16(3):273–280.

Gilleland, E. and Katz, R. W. (2016). extRemes 2.0: An Extreme Value Analysis Package in R. *Journal of Statistical Software*, 72(8).

Gillespie-Marthaler, L., Nelson, K., Baroud, H., and Abkowitz, M. (2019). Selecting Indicators for Assessing Community Sustainable Resilience. *Risk Analysis*, 39(11).

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4(1):26–42.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48(1):26–34. Place: US Publisher: American Psychological Association.

Grazzini, J., Richiardi, M. G., and Tsionas, M. (2017). Bayesian estimation of agent-based models. *Journal of Economic Dynamics and Control*, 77:26–47.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S. K., Huse, G., Huth, A., Jepsen, J. U., Jørgensen, C., Mooij, W. M., Müller, B., Pe'er, G., Piou, C., Railsback, S. F., Robbins, A. M., Robbins, M. M., Rossmanith, E., Rüger, N., Strand, E., Souissi, S., Stillman, R. A., Vabø, R., Visser, U., and DeAngelis, D. L. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, 198(1):115–126.

Gunnthorsdottir, A., Houser, D., and McCabe, K. (2007). Disposition, history and contributions in public goods experiments. *Journal of Economic Behavior & Organization*, 62(2):304–315.

Haimes, Y. Y., Horowitz, B. M., Lambert, J. H., Santos, J. R., Lian, C., and Crowther, K. G. (2005). Inoperability Input-Output Model for Interdependent Infrastructure Sectors. I: Theory and Methodology. *Journal of Infrastructure Systems*, 11(2):67–79.

Hardin, G. (1968). The Tragedy of the Commons. *Science*, 162(3859):1243–1248.

Hasson, R., Lofgren, A., and Visser, M. (2010). Climate change in a public goods game: Investment decision in mitigation versus adaptation. *Ecological Economics*, 70(2):331–338.

He, X. and Liu, H. X. (2012). Modeling the day-to-day traffic evolution process after an unexpected network disruption. *Transportation Research Part B: Methodological*, 46(1):50–71.

Hilbig, B. E., Zettler, I., Leist, F., and Heydasch, T. (2013). It takes two: Honesty–humility and agreeableness differentially predict active versus reactive cooperation. *Personality and Individual Differences*, 54(5):598–603. Place: Netherlands Publisher: Elsevier Science.

Hill, J., Linero, A., and Murray, J. (2020). Bayesian Additive Regression Trees: A Review and Look Forward. *Annual Review of Statistics and Its Application*, 7(1):annurev–statistics–031219–041110.

Hirsh, J. B. and Peterson, J. B. (2009). Extraversion, neuroticism, and the prisoner's dilemma. *Personality and Individual Differences*, 46(2):254–256.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2):179–185.

Horowitz, K. and Planting, M. (2009). Concepts and Methods of the U.S. Input-Output Accounts. Technical report, U.S. Bureau of Economic Analysis.

HVRI (2016). The SoVI Recipe.

Isaak, J. and Hanna, M. J. (2018). User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection. *Computer*, 51(8):56–59.

Isard, W., editor (1998). *Methods of interregional and regional analysis / Walter Isard ... [et al.]*. Number Accessed from https://nla.gov.au/nla.cat-vn1391519 in Regional science studies series (Aldershot, England). Ashgate, Aldershot, Hants, England ; Brookfield, Vt.

Jackson, J., Rand, D. G., Lewis, K., Norton, M. I., and Gray, K. (2016). Agent Based Modeling: A Guide for Social Psychologists. SSRN Scholarly Paper ID 2880587, Social Science Research Network, Rochester, NY.

Janssen, M. and Ostrom, E. (2006). Empirically Based, Agent-based models. *Ecology and Society*, 11(2).

Joerin, J. and Shaw, R. (2011). Chapter 3 Mapping Climate and Disaster Resilience in Cities. In Shaw, R. and Sharma, A., editors, *Climate and Disaster Resilience in Cities*, number 6 in Community, Environment and Disaster Risk Management, pages 47–61. Emerald Group Publishing Limited, Bingley, UK.

John, O. P. and Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In *Handbook of personality: Theory and research, 2nd ed*, pages 102–138. Guilford Press, New York, NY, US.

Johnson, P. M., Brady, C. E., Philip, C., Baroud, H., Camp, J. V., and Abkowitz, M. (2020). A Factor Analysis Approach Toward Reconciling Community Vulnerability and Resilience Indices for Natural Hazards. *Risk Analysis*, page risa.13508.

Johnson, S. (2015). Escaping the tragedy of the commons through targeted punishment. *Royal Society Open Science*, 2(8):150223.

Jumadi, Heppenstall, A. J., Malleson, N. S., Carver, S. J., Quincey, D. J., and Manville, V. R. (2018). Modelling Individual Evacuation Decisions during Natural Disasters: A Case Study of Volcanic Crisis in Merapi, Indonesia. *Geosciences*, 8(6):196.

Kagel, J. and McGee, P. (2014). Personality and cooperation in finitely repeated prisoner's dilemma games. *Economics Letters*, 124(2):274–277.

Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263–292.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23(3):187–200.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20:141–151.

Kapelner, A. and Bleich, J. (2016). bartMachine: Machine Learning with Bayesian Additive Regression Trees. *Journal of Statistical Software*, 70(1):1–40.

Keim, M. E. and Noji, E. (2011). Emergent use of social media: a new age of opportunity for disaster resilience. *American Journal of Disaster Medicine*, 6(1):47–54.

Kline, R., Bankert, A., Levitan, L., and Kraft, P. (2019). Personality and Prosocial Behavior: A Multilevel Meta-Analysis. *Political Science Research and Methods*, 7(1):125–142.

Koo, T. K. and Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2):155–163.

Koole, S. L., Jager, W., van den Berg, A. E., Vlek, C. A. J., and Hofstee, W. K. B. (2001). On the Social Nature of Personality: Effects of Extraversion, Agreeableness, and Feedback about Collective Resource Use on Cooperation in a Resource Dilemma. *Personality and Social Psychology Bulletin*, 27(3):289–301.

Kosinski, M., Stillwell, D., and Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15):5802–5805.

Kotzee, I. and Reyers, B. (2016). Piloting a social-ecological index for measuring flood resilience: A composite index approach. *Ecological Indicators*, 60:45–53.

Kumar, M. and Dutt, V. (2015). Understanding Cooperative Behavior against Climate Change through a Public - Goods Game. *Climate Change*, page 5.

Kunreuther, H. and Michel-Kerjan, E. (2015). Demand for fixed-price multi-year contracts: Experimental evidence from insurance decisions. *Journal of Risk and Uncertainty*, 51(2):171–194.

Lee, Y. S., Seo, Y. W., and Siemsen, E. (2018). Running Behavioral Operations Experiments Using Amazon's Mechanical Turk. *Production and Operations Management*, 27(5):973–989.

Leontief, W. W. (1936). Quantitative Input and Output Relations in the Economic Systems of the United States. *The Review of Economics and Statistics*, 18(3):105.

Levati, M. V., Morone, A., and Fiore, A. (2009). Voluntary contributions with imperfect information: An experimental study. *Public Choice*, 138(1-2):199–216.

Lönnqvist, J.-E., Verkasalo, M., and Walkowitz, G. (2011). It pays to pay – Big Five personality influences on co-operative behaviour in an incentivized and hypothetical prisoner's dilemma game. *Personality and Individual Differences*, 50(2):300–304.

MacKenzie, C. A., Barker, K., and Grant, F. H. (2012). Evaluating the Consequences of an Inland Waterway Port Closure With a Dynamic Multiregional Interdependence Model. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 42(2):359–370.

Maddux, J. E. and Rogers, R. W. (1983). Protection motivation and self-efficacy: A revised theory of fear appeals and attitude change. *Journal of Experimental Social Psychology*, 19(5):469–479.

Mahmassani, H. S. (1990). Dynamic models of commuter behavior: Experimental investigation and application to the analysis of planned traffic disruptions. *Transportation Research Part A: General*, 24(6):465–484.

Marengo, D., Davis, K. L., Gradwohl, G. O., and Montag, C. (2021). A meta-analysis on individual differences in primary emotional systems and Big Five personality traits. *Scientific Reports*, 11(1):7453.

Massey, D., Ahlqvist, O., Vatev, K., and Rush, J. (2018). A Massively Multi-user Online Game Framework for Agent-Based Spatial Simulation. In *CyberGIS for Geospatial Discovery and Innovation*, GeoJournal Library, pages 213–224. Springer, Dordrecht.

McCrae, R. R. and John, O. P. (1987). An Introduction to the Five-Factor Model and Its Applications. *Journal of Personality*, 60(2):175–215.

Mohammad, E. and El-adaway, I. (2017). Sustainable Disaster Recovery Decision-Making Support Tool: Integrating Economic Vulnerability into the Objective Functions of the Associated Stakeholders. *Journal of Management in Engineering*, 33(2):04016041.

Montag, C. and Davis, K. L. (2018). Affective Neuroscience Theory and Personality: An Update. *Personality Neuroscience*, 1:e12.

Montag, C. and Panksepp, J. (2017). Primary Emotional Systems and Personality: An Evolutionary Perspective. *Frontiers in Psychology*, 8.

Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., Schlüter, M., Schulze, J., Weise, H., and Schwarz, N. (2013). Describing human decisions in agent-based models – ODD + D, an extension of the ODD protocol. *Environmental Modelling & Software*, 48:37–48.

Nay, J. J. and Gilligan, J. M. (2015). Data-driven dynamic decision models. In *2015 Winter Simulation Conference (WSC)*, pages 2752–2763, Huntington Beach, CA, USA. IEEE.

Nay, J. J. and Vorobeychik, Y. (2016). Predicting Human Cooperation. *PLOS ONE*, 11(5):e0155656.

Nelson, K. S., Abkowitz, M. D., and Camp, J. V. (2015). A method for creating high resolution maps of social vulnerability in the context of environmental hazards. *Applied Geography*, 63:89–100.

Niles, M. T., Emery, B. F., Reagan, A. J., Dodds, P. S., and Danforth, C. M. (2019). Social media usage patterns during natural hazards. *PLOS ONE*, 14(2):e0210484.

Norris, F. H., Stevens, S. P., Pfefferbaum, B., Wyche, K. F., and Pfefferbaum, R. L. (2008). Community Resilience as a Metaphor, Theory, Set of Capacities, and Strategy for Disaster Readiness. *American Journal of Community Psychology*, 41(1-2):127–150.

Onen, F. and Bagatur, T. (2017). Prediction of Flood Frequency Factor for Gumbel Distribution Using Regression and GEP Model. *Arabian Journal for Science and Engineering*, 42(9):3895–3906.

Orencio, P. M. and Fujii, M. (2013). A localized disaster-resilience index to assess coastal communities based on an analytic hierarchy process (AHP). *International Journal of Disaster Risk Reduction*, 3:62–75.

Ostrom, E. (1999). Coping With Tragedies of the Commons. *Annual Review of Political Science*, 2(1):493–535.

Ostrom, E. (2000). Collective Action and the Evolution of Social Norms. *Journal of Economic Perspectives*, 14(3):137–158.

Paolacci, G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, 5(5):9.

Park, J., Seager, T. P., Rao, P. S. C., Convertino, M., and Linkov, I. (2013). Integrating Risk and Resilience Approaches to Catastrophe Management in Engineering Systems: Perspective. *Risk Analysis*, 33(3):356–367.

Peacock, W., Brody, S., Seitz, W., Merrell, W., Vedlitz, A., Zahran, S., Harriss, R., and Stickney, R. (2010). Advancing Resilience of Coastal Localities: Developing, Implementing, and Sustaining the Use of Coastal Resilience Indicators: A Final Report. Technical report, Hazard Reduction and Recovery Center.

Pendall, R., Foster, K. A., and Cowell, M. (2010). Resilience and regions: building understanding of the metaphor. *Cambridge Journal of Regions, Economy and Society*, 3(1):71–84.

Perfrement, T. and Lloyd, T. (2015). The Composite Resilience Index.

Pfefferbaum, R. L., Pfefferbaum, B., Nitiéma, P., Houston, J. B., and Van Horn, R. L. (2015). Assessing Community Resilience: An Application of the Expanded CART Survey Instrument With Affiliated Volunteer Responders. *American Behavioral Scientist*, 59(2):181–199.

Philip, C. E. and Johnson, P. (2018). The Maritime Safety Journey: An unlikely and remarkable story. *Marine News*, 29(5):48–53.

Poteete, A., Janssen, M., and Ostrom, E. (2010). *Working together: Collective action, the commons, and multiple methods in practice*. Princeton University Press.

Pothos, E. M., Perry, G., Corr, P. J., Matthew, M. R., and Busemeyer, J. R. (2011). Understanding cooperation in the Prisoner's Dilemma game. *Personality and Individual Differences*, 51(3):210–215.

Pregnolato, M., Ford, A., Wilkinson, S. M., and Dawson, R. J. (2017). The impact of flooding on road transport: A depth-disruption function. *Transportation Research Part D: Transport and Environment*, 55:67–81.

Pritchard, J. K., Seielstad, M. T., Perez-Lezaun, A., and Feldman, M. W. (1999). Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Molecular Biology and Evolution*, 16(12):1791–1798.

Prüser, J. (2019). Forecasting with many predictors using Bayesian additive regression trees. *Journal of Forecasting*, 38(7):621–631.

Rachunok, B. A., Bennett, J. B., and Nateghi, R. (2019). Twitter and Disasters: A Social Resilience Fingerprint. *IEEE Access*, 7:58495–58506.

Reilly, A. C., Samuel, A., and Guikema, S. D. (2015). "Gaming the System": Decision Making by Interdependent Critical Infrastructure. *Decision Analysis*, 12(4):155–172.

Renschler, C., Fraizer, A. E., Arendt, L. A., Cimellaro, G. P., Reinhorn, A. M., and Bruneau, M. (2010). Framework for Defining and Measuring Resilience at the Community Scale: The PEOPLES Resilience Framework. Technical Report GCR 10-930, U.S. Department of Commerce National Institute of Standards and Technology, Gaithersburg, Maryland.

Rentfrow, P. J., Gosling, S. D., and Potter, J. (2008). A Theory of the Emergence, Persistence, and Expression of Geographic Variation in Psychological Characteristics. *Perspectives on Psychological Science*, 3(5):339–369.

Revelle, W. (2018). The Personality Project.

Ritchie, L. A. and Gill, D. A. (2007). Social Capital Theory as an Integrating Theoretical Framework in Technological Disaster Research. *Sociological Spectrum*, 27(1):103–129.

Rowcliffe, P., Lewis, M., and Port, A. (1999). The Community Resilience Manual: A Resource for Rural Recovery & Renewal | The Canadian CED Network.

Rubas, D., Hill, H., and Mjelde, J. (2006). Economics and climate applications: exploring the frontier. *Climate Research*, 33:43–54.

Rustichini, A., DeYoung, C. G., Anderson, J. E., and Burks, S. V. (2016). Toward the integration of personality theory and decision theory in explaining economic behavior: An experimental investigation. *Journal of Behavioral and Experimental Economics*, 64:122–137.

Santos, J. R. (2006). Inoperability input-output modeling of disruptions to interdependent economic systems. *Systems Engineering*, 9(1):20–34.

Santos, J. R. and Haimes, Y. Y. (2004). Modeling the Demand Reduction Input-Output (I-O) Inoperability Due to Terrorism of Interconnected Infrastructures*. *Risk Analysis*, 24(6):1437–1451.

Schindler, J. (2012). Rethinking the Tragedy of the Commons: The Integration of Socio-Psychological Dispositions. *Journal of Artificial Societies and Social Simulation*, 15(1):4.

Schroeder, K. B., Nettle, D., and McElreath, R. (2015). Interactions between personality and institutions in cooperative behaviour in humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1683).

Schröder, D. and Gilboa Freedman, G. (2020). Decision making under uncertainty: the relation between economic preferences and psychological personality traits. *Theory and Decision*, 89(1):61–83.

Schweighofer, J. (2014). The impact of extreme weather and climate change on inland waterway transport. *Natural Hazards*, 72(1):23–40.

Seaberg, D., Devine, L., and Zhuang, J. (2017). A review of game theory applications in natural disaster management research. *Natural Hazards*, 89(3):1461–1483.

Shabrina, F. Z., Meilano, I., Windupranata, W., and Hanifa, N. R. (2018). Measure coastal disaster resilience using community disaster resilience index (CDRI) in Mentawai Island, Indonesia. *AIP Conference Proceedings*, 1987(1):020080.

Sharifi, A. (2016). A critical review of selected tools for assessing community resilience. *Ecological Indicators*, 69:629–647.

Shaw, R., Takeuchi, Y., Joerin, J., Fernandez, G., Tjandradewi, B., Chosadilla, Wataya, E., McDonald, B., Fukui, R., Sharma, A., Tsunozaki, E., and Matsuoka, Y. (2010). Climate and disaster resilience initiative capacity-building program - UNISDR.

Sheehan, K. and Pittman, M. (2016). *Amazon's Mechanical Turk for Academics :: The HIT Handbook for Social Science Research*. Melvin & Leigh, Publishers, Irvine, CA.

Sherrieb, K., Norris, F. H., and Galea, S. (2010). Measuring Capacities for Community Resilience. *Social Indicators Research*, 99(2):227–247.

Sisson, S. A., Fan, Y., and Tanaka, M. M. (2007). Sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 104(6):1760–1765.

Solomon, O., Prince, O., and Geosurveys, J. (2013). Flood Frequency Analysis of Osse River Using Gumbel's Distribution. *Civil and Environmental Research*, page 6.

Sorg, L., Medina, N., Feldmeyer, D., Sanchez, A., Vojinovic, Z., Birkmann, J., and Marchese, A. (2018). Capturing the multifaceted phenomena of socioeconomic vulnerability. *Natural Hazards*, 92(1):257–282.

Soto, C. J. and John, O. P. (2017a). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, 113(1):117–143.

Soto, C. J. and John, O. P. (2017b). Short and extra-short forms of the Big Five Inventory–2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality*, 68:69–81.

Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology*, 15(2):201–292.

Stefanelli, A. and Seidl, R. (2017). Opinion Communication on Contested Topics: How Empirics and Arguments can Improve Social Simulation. *Journal of Artificial Societies and Social Simulation*, 20(4).

Swaggart, B. (2021). Port of Cates Landing is America's Newest Multimodal Inland Port - Northwest Tennessee Economic Development.

Tapia, C., Abajo, B., Feliu, E., Mendizabal, M., Martinez, J. A., Fernández, J. G., Laburu, T., and Lejarazu, A. (2017). Profiling urban vulnerabilities to climate change: An indicator-based vulnerability assessment for European cities. *Ecological Indicators*, 78:142–155.

Tavaré, S., Balding, D. J., Griffiths, R. C., and Donnelly, P. (1997). Inferring Coalescence Times From DNA Sequence Data. *Genetics*, 145(2):505–518.

Taylor, R., Forrester, J., Pedoth, L., and Matin, N. (2014). Methods for Integrative Research on Community Resilience to Multiple Hazards, with Examples from Italy and England. *Procedia Economics and Finance*, 18:255–262.

Teyssier, S. (2012). Inequity and risk aversion in sequential public good games. *Public Choice*, 151(1):91–119.

Thiele, J. C., Kurth, W., and Grimm, V. (2014). Facilitating Parameter Estimation and Sensitivity Analysis of Agent-Based Models: A Cookbook Using NetLogo and 'R'. *Journal of Artificial Societies and Social Simulation*, 17(3):11.

THRIVE (2004). THRIVE: Tool for Health & Resilience In Vulnerable Environments | Prevention Institute.

Thurstone, L. (1947). *Multiple factor analysis*. Multiple factor analysis. Chicago, University of Chicago Press.

Toni, T., Welch, D., Strelkowa, N., Ipsen, A., and Stumpf, M. P. H. (2009). Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of The Royal Society Interface*, 6(31):187–202. arXiv: 0901.1925.

Tversky, A. and Kahneman, D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323.

Twigg, J. (2007). Characteristics of a Disaster-resilient Community A Guidance Note. Technical report, DFID Disaster Risk Reduction Interagency Coordination Group.

UNDP (2011). Towards Human Resilience: Sustaining Mdg Progress in an Age of Economic Uncertainty. Technical report, United Nations Development Programme, New York, NY.

UN/ISDR (2005). Hyogo Framework for Action (HFA).

US BEA (2017). Input-Output Accounts Data.

US Census Bureau (2017). The 2017 Commodity Flow Survey Final Tables. Section: Government.

USACE (2017). State to state and region to region, commodity tonnages, public domain database - IWR Reports.

USACE (2021). Rivergages.com: Providing River Gage Data for Rivers, Streams and Tributaries.

USCG (2020). Mississippi River and Tributaries Waterways Action Plan: Upper Mississippi River Annex. Technical report, United States Coast Guard.

Volk, S., Thöni, C., and Ruigrok, W. (2011). Personality, personal values and cooperation preferences in public goods games: A longitudinal study. *Personality and Individual Differences*, 50(6):810–815.

Volk, S., Thöni, C., and Ruigrok, W. (2012). Temporal stability and psychological foundations of cooperation preferences. *Journal of Economic Behavior & Organization*, 81(2):664–676.

Wang, B. and Zhuang, J. (2017). Crisis information distribution on Twitter: a content analysis of tweets during Hurricane Sandy. *Natural Hazards*, 89(1):161–181.

Whalen, L. and Painter, S. (2010). Wal-Mart sets late-delivery fee.

Wilhelmi, O. V. and Hayden, M. H. (2010). Connecting people and place: a new framework for reducing urban vulnerability to extreme heat. *Environmental Research Letters*, 5(1):014021.

Woodruff, J. M. (2005). Consequence and likelihood in risk estimation: A matter of balance in UK health and safety risk assessment practice. *Safety Science*, 43(5-6):345–353.

Wunder, M., Suri, S., and Watts, D. J. (2013). Empirical agent based models of cooperation in public goods games. In *Proceedings of the fourteenth ACM conference on Electronic commerce - EC '13*, page 891, Philadelphia, Pennsylvania, USA. ACM Press.

Yoon, D. K., Kang, J. E., and Brody, S. D. (2016). A measurement of community disaster resilience in Korea. *Journal of Environmental Planning and Management*, 59(3):436–460.

Zhao, K. and Smillie, L. D. (2015). The Role of Interpersonal Traits in Social Decision Making: Exploring Sources of Behavioral Heterogeneity in Economic Games. *Personality and Social Psychology Review*, 19(3):277–302.

Zheng, X. and Cheng, Y. (2011). Modeling cooperative and competitive behaviors in emergency evacuation: A game-theoretical approach. *Computers & Mathematics with Applications*, 62(12):4627–4634.

Zhuang, J. and Bier, V. M. (2007). Balancing Terrorism and Natural Disasters—Defensive Strategy with Endogenous Attacker Effort. *Operations Research*, 55(5):976–991.

Zou, L., Lam, N. S. N., Shams, S., Cai, H., Meyer, M. A., Yang, S., Lee, K., Park, S.-J., and Reams, M. A. (2019). Social and geographical disparities in Twitter use during Hurricane Harvey. *International Journal of Digital Earth*, 12(11):1300–1318.