12-2019

Improving Earthquake Disaster Models with Post-Event Data: Insights from the 2015 Gorkha, Nepal Earthquake

Bradley Steven Wilson
University of Arkansas, Fayetteville

Improving Earthquake Disaster Models with Post-Event Data: Insights from the 2015 Gorkha, Nepal Earthquake

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Geosciences

by

Bradley Steven Wilson
Colorado School of Mines
Bachelor of Science in Geophysical Engineering, 2015
University of Arkansas
Master of Science in Geography, 2016

December 2019
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council

Thomas R. Paradise, Ph.D.
Dissertation Director

Fiona Davidson, Ph.D.
Committee member

Geoboo Song, Ph.D.
Committee member

Jackson Cothren, Ph.D.
Committee member

ABSTRACT

Immense amounts of data are collected following earthquake disasters. Yet, it remains unclear how researchers' might take full advantage of diverse post-disaster datasets. Using data from the 2015 Gorkha Nepal earthquake, this dissertation explores three ways in which post-disaster survey and assessment datasets can be used to inform models of seismic risk, vulnerability, and recovery processes. The first article presents an empirical analysis of scale issues in disaster vulnerability indices using a novel dataset of 750,000 households. This study finds that using aggregated household data to create social vulnerability indices can produce results that are meaningfully different from equivalent indices produced directly with household-level data. These results inform future development of vulnerability indices. The second article develops a Bayesian item-response theory modeling framework for estimating household-level reconstruction behavior from reconstruction progress surveys. This study provides a new way to quantitatively assess earthquake recovery, with results showing large differences in reconstruction probabilities among different levels of aid receipt, household willingness to commit additional resources, and geographic location. The final article uses engineering damage assessment data to develop a model for spatially interpolating geolocated clusters of rapid damage assessments onto a high-resolution grid. Incorporating ground truthed data significantly improves existing rapid estimates for completely damaged buildings and is feasible with the current scope of rapid damage assessment collection. Together, these contributions cast a vision for an improved disaster modeling ecosystem that more effectively integrates novel post-disaster data streams.

Keywords: Disaster Risk Reduction, Vulnerability, Reconstruction, Earthquakes, Nepal

ACKNOWLEDGEMENTS

This dissertation is the result of many afternoon runs, too many cups of coffee to count, and the gentle unspoken encouragement of those around me. There were, of course, many more cumulative hours of writing, frustration, and late nights fixing figure captions, but somehow those tend to blur together and fade away. It's a bit surreal to finally stare at a completed document—three years of life collapsed into some 100 odd pages. If I've learned anything over the past few years though, it's that the words on the page only tell part of the story. While the chapters might seem like they're about statistical models and disaster risk reduction, they're really about adventures in Nepal, camping trips with friends, and happy hour conversations. At this point, any attempt to separate the two seems futile.

I'm endlessly thankful for the unwavering support of most everyone along this journey. Sparing pages and pages of gratitude, these are a few of the names most important to me during the past few years: Tom, Fiona, Jack, Geoboo, Aaron, Natalie, Steve, Rebecca, Alex, Emily, Luke, Elliot, Austin, Brendan, Rayna, Karuna, Lauren, Karuna, Megan, Paulo, Nicole, Jon, Nama, Gaurav, Kate, David, and Eric. Additional thanks to Kathmandu Living Labs for hosting me for five months in Nepal and to the Sturgis family for providing financial support for for my field work.

This research was performed under a National Science Foundation Graduate Research Fellowship, Grant: DGE-1450079. Additional research funding was provided by the University of Arkansas.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PUBLICATIONS

Chapter 2: Wilson, B.S. (2019). Overrun by averages: an empirical analysis into the consistency of social-vulnerability components across multiple scales. *International Journal of Disaster Risk Reduction*, 40.

Chapter 3: Wilson, B.S. (2019). A Bayesian modeling approach for estimating earthquake reconstruction behavior. *Annals of the American Association of Geographers*, (In Review).

Chapter 4: Wilson, B.S. (2019), A spatial interpolation model for high-resolution mapping of earthquake damages from geolocated cluster data, (In Preparation).

1: INTRODUCTION

1.1 Motivation

Statistical models of earthquake outcomes are crucial elements of disaster risk reduction (DRR) and management in seismically active areas. Spanning a wide range of subject matter from hazards to vulnerabilities, models formalize existing knowledge on the factors affecting earthquake impacts in order to mitigate losses in future events. One of the longstanding challenges in earthquake modeling efforts is the effective organization and synthesis of relatively sparse data sets (Kessler and Hendrix, 2009, Zhang et al., 2015, Li et al., 2019). Given that earthquakes occur infrequently, the number of relevant datasets in any given area has historically been limited. As a result, the types of approaches typically seen in DRR studies—vulnerability indices, perception surveys, and empirical damage models to name a few—are tailored to work with data that is consistently available across multiple events. While favoring incremental development and validation of existing models is not problematic in itself, there is a significant lack of research analyzing how new data streams should be integrated into existing modeling efforts.

In recent years, the amounts of data collected in the wake of earthquake disasters has skyrocketed on the backs of crowd-sourcing efforts, data-driven development aid initiatives, and the broader open data movement (Amin and Goldstein, 2008, Goodchild and Glennon, 2010, Li et al., 2019). Furthermore, these new streams of post-disaster data are increasingly made openly available through data repositories like the Humanitarian Data Exchange and ReliefWeb. In theory, a robust set of open data provides the necessary platform for researchers to innovate on disaster models. However, coordination among data-producers is limited, leading to a collection of heterogeneous, geographically distributed, and thematically variable datasets with multiple formats and standards. Initiatives like the Humanitarian Exchange Language" (Kessler and Hendrix, 2009) are attempting to provide some standardization and interoperability of various datasets by providing common data 'tags', but adoption is limited and the tags only benefit certain types of datasets. As it currently stands, the open-data movement offers a promising vision for future DRR research, but the roadmap for implementation remains unclear (Li et al., 2019).

This dissertation targets this shortcoming by providing three clear examples of how different types of post-disaster data can inform statistical models across the DRR spectrum, including social vulnerability assessment, recovery monitoring, and rapid impact evaluation. All three chapters focus on the 2015 Gorkha, Nepal Earthquake, an event that exemplifies the new types of data streams that are becoming openly available. The philosophy underlying this dissertation is one of supporting the innovation or adaption of methods to handle new data rather than attempting to adapt specific datasets to work with existing methods. The contributions are in no way intended to supplant existing methods or stand alone as a complete framework, but rather to jointly and independently offer a promising new set of pathways for disaster modeling. For this reason, the thematic scope of the dissertation is kept intentionally broad, mirroring one of the four main priorities of the 2015 Sendai Framework for Disaster Risk Reduction in "Understanding all areas of disaster risk, including "vulnerability, capacity, exposure of persons and assets, hazard characteristics and the environment," (UNDRR, 2015).

1.2 Scope and organization of contributions

This dissertation covers three distinct concentrations within disaster risk reduction: social vulnerability, reconstruction, and rapid impact assessment. However, all three chapters have the unified goal of using novel post-disaster data sets to improve the types of statistical models used in disaster risk reduction research. To increase cohesion between the different models and topics discussed in each chapter, all three studies utilize data from the 2015 Gorkha, Nepal Earthquake. Any background and contextual information presented henceforth is generally applicable to all three studies.

Several guiding principles underlie the model development process used in this dissertation. These principles provide a consistent set of objectives for all chapters and specify the types of model improvements prioritized in this research. The principles are listed as follows:

1. **Prioritize disaggregated results:**

   Earthquake risks and vulnerabilities are often analyzed at aggregated scales. While there certain situations where aggregated statistics may be preferable, more often than not the use of aggregated statistics is simply a matter of data availability. Relying on aggregated data can mask or bias results when the phenomena of interest operates at a finer scale than the unit of analysis. Emergency response activities, mitigation efforts, and reconstruction programs all ideally target specific individuals, households, or communities, but capturing detailed variability at these scales is challenging due to data or resource constraints. All of the studies in this dissertation focus on using large, geographically distributed household-level or equivalent datasets to improve the resolution at which disaster processes are modeled.

2. **Account for spatial variability:**

   Earthquakes affect large and diverse geographic areas. While many elements of emergency response and reconstruction planning are funneled through centralized authorities, most evidence points to strong heterogeneities in disaster impacts. Although many areas of rural Nepal are physiographically similar, there are distinct differences among cultural histories and the impact of dynamic social processes like urbanization and labor migration. Therefore, it is important to consider potential spatial variations among both environmental and social components of disaster processes. All present analyses incorporate spatial effects when relevant, including accounting for spatial autocorrelations in damage levels and modeling geographic variability in disaster aid receipt.

3. **Explicitly address uncertainty:**

   There are numerous sources of uncertainty within earthquake hazard, risk, and vulnerability models. From ground motion measurement to survey collection and construct reliability, disaster risk management is an exercise in decision making under uncertainty. Rather than mask any assumptions in the modeling process, all of the work in this dissertation attempts to directly account for and visualize uncertainties. This goal takes several on different forms across the three chapters, including running simulation studies, comparing multiple hypothesis, and using Bayesian statistical frameworks that include full uncertainty estimates on model parameters. Additionally, several chapters highlight the current state of uncertainty representation (or lack there of) in existing model designs and offer tangible pathways for improvement.

This rest of this report is organized into three primary chapters, each consisting of a self-contained manuscript with appropriate background, analysis, discussions, and conclusions. Chapter one covers social vulnerability and uses household-level socio-economic data to address scale-related issues in index-based methodologies. Chapter two focuses on reconstruction and recovery, developing a model for analyzing reconstruction-related behaviors from household livelihood and perception surveys. Chapter three addresses impact modeling, using damage assessment data to validate an alternative spatial modeling approach for early-estimation of structural damages. Together, the three chapters cast a vision for an improved disaster modeling ecosystem that more effectively integrates novel post-disaster data streams. A brief overview of the Gorkha Earthquake is provided below followed by a summary of motivations and primary research objectives for each chapter.

1.2.1 The 2015 Gorkha, Nepal Earthquake

Situated in the middle of the Himalayan mountain belt, Nepal is one of the most earthquake prone countries in the world (Dangal, 2011), experiencing at least one deadly earthquake per century over Nepal's recorded history. This dissertation focuses on the most recent earthquake, a magnitude 7.8 (M_w) event striking central Nepal on April 25, 2015. Several large aftershocks followed the main rupture including a magnitude 7.3 (M_w) event on May 12 occurring further east of the mainshock. The primary shallow thrust rupture initiated in the district of Gorkha and propagated eastward towards the Kathmandu valley, causing extensive damage across Nepal's hill and mountain districts. According to official statistics, the Gorkha earthquake resulted in 8686 deaths, 22300 injuries, and left 2.2 million people homeless (Government of Nepal, 2015). The earthquake also significantly damaged infrastructure, fully destroying over 500,000 houses and partially affecting another 250,000+ (Government of Nepal, 2015). Total estimated economic damages are approximately 10 billion USD, roughly 50% of Nepal's gross domestic product (Joshi and Joshi, 2018).

Thirty-one of Nepals 75 districts were designated as earthquake-affected by the Nepali government. Of the 31 districts, fourteen were designated as highly affected: eleven rural districts and the three districts comprising the Kathmandu Valley. The highly affected districts comprise the study region for all three chapters, with two chapters focusing just on the eleven rural districts. Most of the damage was spread across Nepal's hill and mountain regions, where traditionally constructed stone and mud structures account for 50% of all buildings, with totals reaching closer to 80%-90% in rural areas. Field reconnaissance surveys found buildings constructed on ridges to be more severely affected than those constructed on shallow slopes (Parajuli and Kyono, 2015). Contrary to expectations, most of the structures in the Kathmandu Valley performed relatively well during the earthquake due to a lack of resonance effects between the dominate low frequency ground motions and the high natural frequencies of the predominate building types in the area (Parajuli and Kyono, 2015, Rai et al., 2016). Damage patterns in the valley were highly clustered, suggesting potential site amplification or localized failures to enforce seismic building codes (Goda et al., 2015).

Nepal's recovery from the Gorkha Earthquake has been slow and contentious. The first year of post-disaster relief and recovery activities was situated against a contentious political background that delayed recovery activities (Comfort and Joshi, 2017). The National Reconstruction Authority (NRA)—the governing body designated to oversee the reconstruction process—was established by political ordinance in late June of 2015, but the associated bill failed to pass Parliament in August and the NRA was subsequently dissolved. It took several months before the NRA was reestablished on December 25, 2015, eight months after the earthquake. This intermittent period was characterized by political unrest and transition related to Nepal's proposed new Constitution that outlined a new federal structure and administrative boundaries. The Constitution was formally approved in September despite fierce protests and boundary disputes from the Madhesi people, one of Nepal's largest ethnic groups living primarily in the Terai Region bordering India. Imports to Nepal from India were dramatically reduced during this period in an alleged unofficial Indian blockade. Border issues and extreme fuel shortages continued through March of 2016, severely impacting the Nepali economy and further delaying reconstruction activities. In May of 2016, the Post Disaster Recovery Framework was published and recovery work formally started (Government of Nepal, 2016).

1.2.2 Chapter 2: Addressing scale issues in social vulnerability indices

Social science research has uncovered many ways in which individuals, communities, and societies are differentially impacted by earthquake events. These differences, broadly termed vulnerabilities if positively associated with more severe impacts or coping capacities if negatively associated with impacts, focus not only on physically unsafe conditions, but also on the larger systems, policies, and social conditions that evolve such conditions in the first place (Cutter et al., 2003, Wisner et al., 2004, Thomas et al., 2013). As a result, the scope of vulnerability research is quite broad—a feature that has supported multiple divergent research directions but limited clear consensus building.

The alignment of theory with empirical evidence is a longstanding issue in social vulnerability modeling. Social vulnerability research has its roots in qualitative, localized case studies that sought to understand the context-specific factors contributing to adverse disaster impacts (Cutter et al., 2003, Schmidtlein et al., 2008). A focus on people—individuals, households, and communities—is a central theme in seminal vulnerability texts. However, the past several decades have seen marked increases in the number of national or sub-national scale social vulnerability indices (Khazai et al., 2014, Beccari, 2016). In theory, areal indices offer generalized, spatially comparable metrics that can meaningfully supplement physical hazard models (Tate, 2012, 2013). However, it is not currently clear whether the characteristics derived from local qualitative research generalize well to aggregated scales.

The first chapter of this dissertation contributes to this alignment discussion by providing an empirical analysis of the consistency of social vulnerability components between household and aggregated scales. Somewhat surprisingly, there is no consistent rationale for variable inclusion in social vulnerability indices and the assumptions underlying variable selection are rarely articulated (Beccari, 2016). Many vulnerability indices include aggregated measures of household characteristics on the basis that it seems reasonable to assume that factors relevant to household vulnerabilities (e.g. age, gender, poverty, disability status) are equally relevant in aggregate (Schmidtlein et al., 2008). This study challenges this assumption, using a complete set of household level micro-data (750,000+ households) collected after the Gorkha Earthquake to show that social vulnerability dimensions estimated from household-level data are different from those estimated at areal scales—using otherwise identical data. The implications of these qualitative and quantitative differences between household and aggregated scales are discussed both in the context of Nepal and for social vulnerability indices at large.

1.2.3 Chapter 3: Developing a modeling framework for household reconstruction behavior

In recent earthquakes, reconstruction frameworks have increasingly shifted towards favoring 'owner-driven' approaches where greater decision making authority is given to local actors (World Bank, 2015). Macro-level reconstruction standards and policies are still organized by a centralized authority, but individual households are largely free to make their own reconstruction decisions. Nepal's reconstruction framework is no exception, explicitly incorporating many principles of 'owner-driven' reconstruction (Government of Nepal, 2016). One of the key elements of the framework is the Rural Housing Reconstruction and Recovery Program (HRRP), a multi-phase, multi-stakeholder project aimed at providing the support necessary to guide the 'owner-driven' reconstruction process. Eligible households are entitled to financial support in the form of a NPR 300,000 three tranche grant (~$3000 USD) and other social and technical support mechanisms contingent upon certain stipulations including payment structure, construction standards, and grievance redress procedures.

Owner-driven reconstruction frameworks are motivated by a 'local solutions to local problems' mindset and the notion that providing local communities ownership over the reconstruction process leads to stronger recovery outcomes (Mishra et al., 2017). Despite these intentions, perceptions vary over the extent to which local realities have matched these goals (Daly et al., 2017, Hall et al., 2017, Mishra et al., 2017, Bownas and Bishokarma, 2018, He et al., 2018). In Nepal, only a fraction of municipalities have received the full range of social and technical assistance, leading to key shortages in the provision of engineering consultations necessary to verify reconstruction progress. The grant dispersal agreement requires that certain stages of reconstruction progress be verified as adhering either to specific building designs or to the minimum standards of Nepal's National Building Code prior to disbursement of the 2nd and 3rd tranches. As a result, household decisions to start the reconstruction process are technically 'owner-driven', but potentially contingent upon certain types of aid receipt.

While aid distribution statistics are tracked, no previous research has attempted to model how aid receipt and other factors contribute to actual household-level decision making. Building upon item-response theory models from health geography that use survey responses to estimates of behavioral action (Alegana et al., 2017, 2018), this chapter develops a Bayesian modeling framework for understanding household-level reconstruction behavior. The key contribution lies in linking estimates of latent household-level reconstruction ability to survey item parameters that allow reconstruction probabilities to vary geographically. Just under 6,000 responses from post-earthquake reconstruction perception and livelihood needs surveys are used to fit the item-response theory model. This chapter's findings show the importance of engineer availability and indicate that households that are able and willing to commit their own resources to reconstruction receive significantly larger benefits from government provided assistance compared to other households. Providing robust, quantitative descriptions of these differences is an important contribution for reconstruction progress monitoring and better understanding how to achieve equitable aid distribution in owner-driven systems.

1.2.4 Chapter 4: Improving modeled damage estimates with geolocated cluster data

Accurately determining the severity, extent, and location of severe damage is a key challenge following major earthquake activity (Goodchild and Glennon, 2010, Erdik et al., 2011, Lallemant et al., 2017). Although remotely sensed imagery and eye-witness reports provide some early indication of damage levels, spatially consistent assessments of damage remain limited. Consequently, rapidly modeled damage estimates that rely on ground motion and building fragility estimates often serve as the best available source of event-level impact despite their limitations (Lallemant et al., 2017). These types of damage models are generally accurate for 'order-of-magnitude' estimates, but struggle to capture spatial damage variability, often due to limitations in input data (Erdik et al., 2011, Jaiswal et al., 2011). Averaging shaking intensities and housing data across large areas neglects local ground motion variability and differences in collapse probabilities among different housing typologies.

Improving the rapid damage models is important because their results inform requests for international aid. The 2008 Joint Declaration on Post-Crisis Assessments and Recovery Planning, signed by the European Commission, United Nations, and World Bank, details a common process by which disaster impacts are evaluated and recovery assistance is mobilized. Central to this process is the creation of a multi-sectoral damage and needs assessments. However, as Lallemant et al. (2017) notes, the timeline required to deliver the Post Disaster Needs Assessment is too short ($\sim$ 1 month) to fully survey event damages and PDNAs often end up based on ad-hoc collection, analysis, and extrapolation of available data. In the weeks following the Gorkha Earthquake, over 60,000 rapid field damage assessments were collected by teams of trained engineers, yet this information was not used in the damage estimates calculated for the PDNA. Instead, each affected district was assigned an shaking intensity and housing counts for four different building typologies and fragility curves were used to estimate the number of partially collapsed and totally collapsed houses per building type (Government of Nepal, 2015). These estimates suggested that upwards of 90% of residential structures were completely destroyed across eleven priority districts. While these figures were met with skepticism, they were considered the 'best-available' event-level statistics available (Lallemant et al., 2017).

The third chapter evaluates an alternative modeling strategy for estimating earthquake damage that takes advantage of rapidly available field damage assessments. This approach leverages recent computational strategies for spatial modeling that would have previously been intractable (Rue et al., 2009, Lindgren et al., 2011). Instead of relying on fragility curves and areal shaking estimates, the proposed approach models the statistical relationship between gridded covariates and ground-truthed damage states at geolocated survey clusters. These correlations are combined with estimates of spatial autocorrelation to interpolate damage and uncertainty estimates onto a uniform prediction grid. Model predictions with varying number of simulated survey clusters are validated against a complete set of damage data collected further along into the reconstruction process. Results show improved predictive capacity with more spatial resolution in comparison to more standard damage estimation procedures—a promising step towards improving damage modeling capacities in the pre-PDNA timeframe.

2: AN EMPIRICAL ANALYSIS INTO THE CONSISTENCY OF SOCIAL-VULNERABILITY COMPONENTS ACROSS MULTIPLE SCALES

This chapter corresponds to the following published paper: Wilson, B.S. (2019), Overrun by averages: an empirical analysis into the consistency of social-vulnerability components across multiple scales, *International Journal of Disaster Risk Reduction*, 40.

2.1 Abstract

Social vulnerability indices have become widely accepted as key elements of disaster risk reduction (DRR) frameworks despite significant conceptual and technical concerns over their creation. Despite increasingly complex theorization, it remains commonplace to determine vulnerability dimensions by applying principal component or factor analysis on aggregated, sub-national level statistical data. The primary focus of this paper is in showing that social vulnerability dimensions derived from aggregate statistics may not be consistent with those derived at the household level. I first provide empirical support for this problem using household level micro-data from Nepal, then illustrate how differences in determined dimensions impact composite index creation at both household and regional levels. I show that qualitative differences in index components are not required to produce large magnitude changes or sign reversals in index scores. These findings raise questions over whether regional vulnerability components can be sufficiently represented with household characteristics. These results reiterate the necessity of addressing issues of scale in the development of empirical social vulnerability indicators.

Keywords: Social vulnerability, Vulnerability indicators, Ecological fallacy, Disaster risk reduction, Nepal Earthquake

2.2 Introduction

Models of social vulnerability are becoming integral components of disaster risk reduction (DRR) frameworks. These models are important for describing the dynamic social, economic, and political factors that contribute to differential disaster impacts (Cutter et al., 2003, Turner et al., 2003, Wisner et al., 2004, Anbarci et al., 2005, Keefer et al., 2011, Neumayer et al., 2014). Social vulnerability helps inform an understanding that people living at the margins are vulnerable not simply because they are more likely to live in substandard housing or are located in a hazardous area, but also because they might lack access to things like social services and political representation. These complex, multi-dimensional factors interact and compound over time to limit the ability of individuals, communities, or societies to prepare for, withstand, and recover from natural disasters.

Empirical vulnerability analyses in particular have increased in prevalence over the past several decades as part of larger, multi-disciplinary efforts to jointly model vulnerability and physical hazards (Khazai et al., 2014, Beccari, 2016). Among empirical studies, the development of vulnerability indices is far and away the most popular approach. Indices rely on various algorithms to manipulate and aggregate demographic data to describe spatial distributions of potential vulnerabilities (e.g. Wu et al., 2002, Cutter et al., 2003, Chakraborty et al., 2005, Rygel et al., 2006, Borden et al., 2007). A recent review found over a hundred published risk, vulnerability, and resilience indices in the DRR and related literature between 1995 and 2015 (Beccari, 2016).

Despite their relative popularity, vulnerability indices have amassed a fair amount of skepticism (Barnett et al., 2008). Several broad theoretical frameworks have emerged (Cutter et al., 2003, Turner et al., 2003, Birkmann, 2007), but a lack of consistency over data sources, methods, and scales have made it difficult to synthesize results into general practice. Validating results has also proved challenging, both internally with respect to index uncertainty and sensitivity (Schmidtlein et al., 2008, Damm, 2010, Tate, 2012, 2013), and externally by comparing results to independently observed measures of harm (Fekete, 2009, Burton, 2010, Schmidtlein et al., 2011). Although having a diversity of indices does theoretically provide some benefit for researchers or end-users looking to tailor an approach to a specific situation, it is not clear if current index methods generalize well.

Scale issues, both conceptual and methodological, contribute to the generalization problem. Although semantics vary, vulnerability almost always is defined as to encompass processes operating at and across multiple scales (Birkmann, 2007, Fekete et al., 2010). For example, Wisner et al. (2004) describes a progression of vulnerability from diffuse 'root causes' operating at institutional or ideological scales to 'unsafe conditions' that manifest at the local scale—all the way down to the individual level. Identifying relevant variables and their interactions is difficult across such a wide range of potential scales, let alone tracing chains of casual explanation (Wisner et al., 2004). As vulnerability theories are transformed into conceptual models or quantified with empirical methods, these complex notions of scale hierarchies are often lost (Fekete et al., 2010). Previous vulnerability research has come a long ways in identifying different qualitative drivers of vulnerabilities, but there remains a wide range of perspectives over how to best transform available empirical data into adequate representations of complex, multidimensional processes.

One of the key methodological challenges in index development is selecting indicator variables to represent underlying vulnerabilities. Early vulnerability indices, including the original Cutter et al. (2003) SoVI, predominately relied on local-scale qualitative research to inform variable inclusion. Many of these household or individual level characteristics are now firmly embedded in empirical vulnerability theory as a result of several decades of incremental developments on top of these early indices Phillips et al. (2013). It remains common to justify particular index variables on the basis of previous inclusion in other indices (Beccari, 2016). While it seems reasonable to assume that some of the latent characteristics relevant to households or individuals are also relevant to communities or societies, there is a notable lack of research evaluating the statistical practices associated with such generalizations. Existing sensitivity and uncertainty analyses (e.g. Jones and Andrey, 2007, Schmidtlein et al., 2008, Tate, 2012, 2013) have examined a variety of index construction steps, including scale specification, but few consider local, household, or individual scales. As a result, the consistency of vulnerability index results between household and other areal scales is still largely unknown.

Using a novel micro-data set from Nepal, this paper aims to empirically evaluate the consistency of vulnerability index results produced at both aggregated and disaggregated scales. This methodological analysis probes into the question of whether vulnerability index results produced with aggregated data adequately represent vulnerabilities of the corresponding households. Focus is placed on variables that are commonly used in both household and regional analyses. Towards this goal, the rest of this paper is organized as follows. Section 2 provides a more detailed overview of aggregation problems and their relevance to modeling vulnerability, both conceptually and methodologically. Section 3 provides a basic context of the study area and additional rationale for the specific choice of micro-data set. In Sections 4 and 5, I outline a general inductive index approach for comparing vulnerability scores between household and regional levels and provide the corresponding results. Section 6 contains a more detailed discussion of results, centering mostly around implications for further aligning empirical vulnerability methods with underlying vulnerability theory—an active discussion in vulnerability literature (Rashed and Weeks, 2003, Birkmann, 2007, Barnett et al., 2008, Tate, 2013).

2.3 Aggregation issues in vulnerability indices

The empirical concerns related to analysis scale are often described as aggregation or modifiable areal unit problems. In areal analyses, the relationships between variables are sensitive to aggregation scale, method, and boundaries (Clark and Avery, 1976, Openshaw, 1983). As a result, variable associations among aggregated populations are not necessarily consistent with those in the corresponding disaggregated subpopulations. Differences can range from small changes in association magnitude to sign reversal in extreme cases—a phenomena known as Simpson's Paradox. Therefore, it is often recommended that results are interpreted only at the scale(s) at which the analysis was performed. Common violations of this principle include the ecological fallacy (Robinson, 1950), where conclusions about individuals are derived from group analyses, and the individualistic fallacy, where conclusions about groups are derived solely from individual analyses (Subramanian et al., 2009). Practically, this means most vulnerability indices are developed at a single scale, despite many formal definitions of vulnerability using scale-invariant language (e.g. 'the element at risk') Beccari (2016).

Although terms like 'ecological fallacy' are often referenced in an empirical or statistical context, they have conceptual corollaries that are more broadly discussed. Birkmann (2007) refers the issue of 'contextualization', or the concern that specific variables can be differentially relevant in different contexts or at different scales. This both describes the case where certain data captures a particular concept more comprehensively in one context versus another and the case where certain indicators are only applicable in a specific context. Fekete et al. (2010) also discusses contextualization in a comparison of vulnerability assessments produced at different scales. In this study, the authors detail the informational and contextual differences between vulnerability assessments at local, sub-national, and national scales, reiterating concerns over generalizability and unclear causal relations between scale levels.

Several uncertainty and sensitivity analyses have examined the effects of analysis scale on areal index results. Tate (2012, 2013) found that social vulnerabilities indices are somewhat sensitive to the choice of scale, although the magnitude of the effect varied by model specification. Among the three most common vulnerability index designs (deductive, inductive, and hierarchical), inductive methods were the most sensitive to changes in analysis scale. Schmidtlein et al. (2008) also included analysis scale in a sensitivity analysis performed on the Social Vulnerability Index (SoVI), a prominent inductive index algorithm. The authors found that the algorithm's results explain less of the total variance in the data as the scale of analysis decreases, but otherwise the qualitative interpretation of results remains similar. On the whole, these studies suggest analysis scale is an important consideration in vulnerability assessments, but the associated sensitivities are relatively modest. Importantly, however, these particular analyses only consider U.S. census block groups and census tracts as potential analysis scales. The differences between aggregated and disaggregated index results remain under-explored.

It has been suggested that an idealized vulnerability assessment would a detailed qualitative identification of vulnerability drivers and selection of variables accordingly (Phillips et al., 2013, Tate, 2013). However, evidence from Beccari (2016) indicates this practice is not the norm. Among 106 reviewed methodologies, variable inclusion in other indices and data availability were common reasons for variable selection. While there is nothing inherently wrong about using similar indicators between studies, it can be problematic to assume that data sources or variables from one study show statistically similar trends in another context. There is a crucial difference between the proxy variables used in vulnerability assessments and the underlying constructs they represent. The data used in vulnerability assessments are indicative rather than absolute representations of constructs. However, few studies consider the measurement reliability of specific data sources, particularly with respect to scale. As Schmidtlein et al. (2008) notes, the relationship between aggregated and disaggregated measures of the same construct are often unknown. Thus, although it seems sensible to assume similar trends across scales, the potential implications of aggregation problems need to be considered—especially given the long-standing theoretical importance of household-level characteristics to vulnerability research.

2.4 Study Area & Data

This study uses micro-data from Nepal's Household Registration for Housing Reconstruction Survey (HRHRP). The HRHRP survey, led by Nepals Central Bureau of Statistics with support from the Ministry of Foreign Affairs and Local Development and National Reconstruction Authority, was designed to collect structural engineering and socio-economic information for households affected by the 2015 Nepal earthquake (also called the Gorkha earthquake) (Ghimire, 2016 (accessed October 7, 2018). The Gorkha earthquake had severe impacts across rural Nepal, destroying over 500,000 buildings, displacing 2.6 million individuals (NSET, 2015). Several post-disaster case studies have pointed to social vulnerabilities as a key factor in understanding the impacts of the Gorkha earthquake and associated recovery processes (DeYoung and Penta, 2017, He et al., 2018, Shapira et al., 2018).

Phase I of the HRHRP survey covers the study region, and was implemented using a census model where the household(s) associated with every residential building in the 11 most severely-affected rural districts (Figure 2.1) were surveyed regardless of the level of earthquake damage or household grant eligibility. The 11 included districts cover all of the rural priority affected districts by the Government of Nepal. Thus, this data provides a unique set of non-sampled household-level data for rural Nepal, regions where mountainous terrain and poor road connectivity have historically made it difficult to collect data at scale. This is not unlike other countries, where detailed socio-demographic data is often limited in scope to a relatively small number of locations. That said, the HRHRP data is somewhat limited in categorical scope compared to more traditional census data. For this study, eleven household-level variables (see Appendix A) are used in the analysis. This is fewer variables compared to other indices, but most common theoretical constructs are represented, including gender, poverty, age, education, family structure, disability status, and migration.

In Nepal, many of the common indicators used in vulnerability assessments are linked to the country's transitional social landscapes (Gentle and Maraseni, 2012, Jaquet et al., 2016, Mainali and Pricope, 2017, He et al., 2018). Nepal is the least urbanized country in South Asia, but is rapidly urbanizing. As a result, both domestic and international migration have become important features of Nepali society. Domestic migration includes both 'push migrants' displaced by natural disasters, civil conflict, lack of job opportunities, or poor public services, and those searching for better economic or educational opportunities (Fafchamps and Shilpi, 2009). International migration is dominated by labor migration (Muzzini and Aparicio, 2013). Nepal has one of the highest labor exportation rates in the world and ranks third globally for remittance income proportion of gross domestic product (30%). From 2001 to 2011 alone, Nepals absentee population increased from 760,000 to 1.9 million individuals, increasing the national absent population percentage to over 7% percent (Muzzini and Aparicio, 2013).

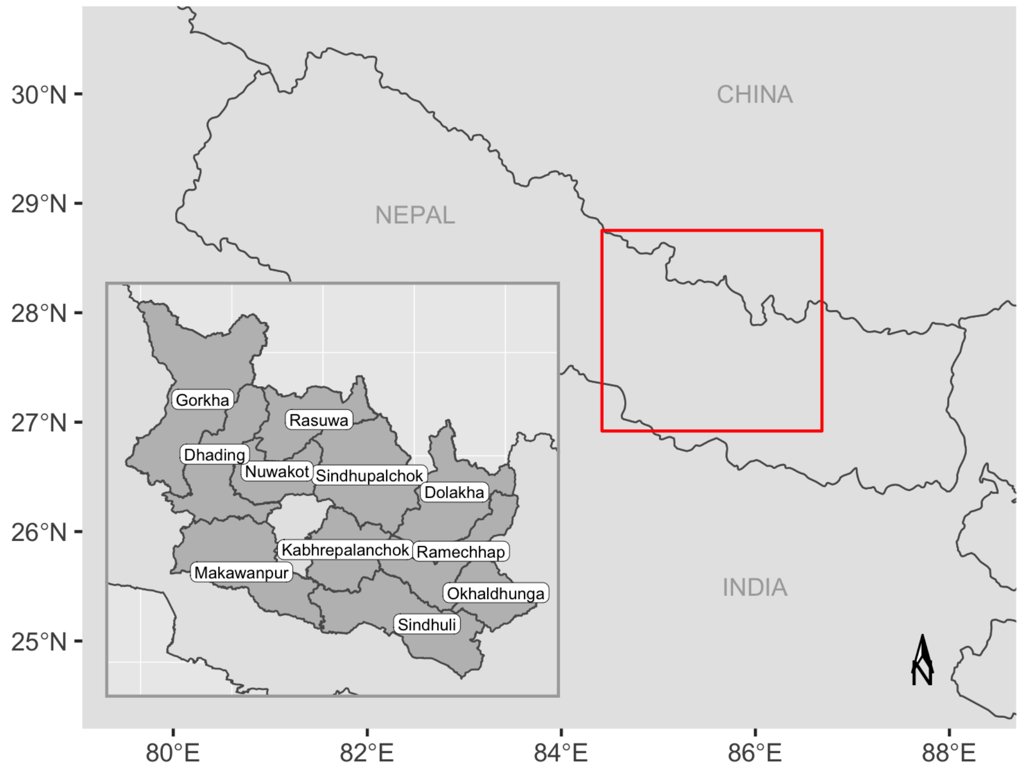

Figure 2.1: Location of the eleven rural districts included in the Nepal's HRHRP Phase I survey dataset. Data for the Kathmandu Valley is not included in the open dataset because of differences in sampling methodology.

Migration has had cascading effects on the demographic and social structures present in many rural communities. Although a ban on women's labor migration was partially lifted in 2002, outmigrants remain overwhelmingly young and male (90%) (Thieme and Wyss, 2005). In addition to skewing demographic distributions, a gender and age-related shift away from agricultural livelihoods towards remittance livelihoods has feminized Nepal's agricultural sector and contributed to increased land abandonment (Gartaula et al., 2012, Tamang et al., 2014, Paudel et al., 2014, Jaquet et al., 2016). This has placed greater responsibility on female and elderly household heads and has been linked to changes in land management and agricultural decision making. There is also some evidence that these types of social changes may exacerbate poverty and contribute to spatial inequalities between sustenance farmers on degraded land and those who have migrated to urban areas (Sunam and McCarthy, 2016). Thus, although fewer total variables are present in the HRHRP micro-data, the available socio-demographic variables (gender, age, migration, family structure, etc.) are likely reflective of the predominate social characteristics in the study area.

The HRHRP micro-data tables were downloaded from the 2015 Nepal Earthquake Open Data Portal ([dataset] Kathmandu Living Labs). These include data for both households and individuals, with the individual records nested in a corresponding household. For this analysis, the individual records were processed to add several household level variables (household size, percent of household physically present, outmigration presence, presence of children/elderly, disability presence). Additionally, each variable was numerically coded and re-leveled such that increasing values are directionally associated with increasing social vulnerability in a matter consistent with previously published associations. This ensures that the individual components are not canceling each other out in composite score calculations. After processing the data and removing incomplete records (.01% of total), the data covers 741,067 households in 110 municipalities across 11 districts. Survey data for the Kathmandu Valley (Kathmandu, Bhaktapur, Lalitpur) is not included in the open-data because of differences in sampling methodology.

2.5 Methods

The overarching goal of this paper is to empirically analyze the underlying aggregation issues present in vulnerability indices. To these ends, this paper uses household level micro-data to evaluate the consistency of social vulnerability index scores estimated using a principal component analysis (PCA) based inductive model structure at both household and municipality levels.

This process involves applying PCA to both the household micro-data and municipality-upscaled equivalents to determine vulnerability components and associated variable loadings. These loading schemes are then used to compare composite vulnerability scores calculated in several ways at both household and municipality levels. An inductive index design was selected because the process of dimension identification step is governed by statistical techniques rather than choices made by the index designer. This decision does not guarantee more accurate or reliable results compared to other models, but it does provide an algorithm that is both applicable and intuitive to use at both household and regional levels. Additionally, previous sensitivity analysis research highlighted inductive approaches as the most sensitive model structure with respect to analysis scale (Tate, 2012), making them a prime candidate for further study. It is worth emphasizing here that the focus of this paper is methodological—it is not intended to be vulnerability assessment. While every attempt has been made to use relevant data and simulate a realistic index design process, the social vulnerability 'scores' presented here are not intended for use in an applied setting.

2.5.1 Comparing inductive component consistency

The inductive approach adopted for this study uses principal component analysis (PCA) to reduce a set of variables into linearly uncorrelated components that serve as proxies for vulnerability dimensions. In general, for a given n-by-k data matrix $\mathbf{X}$ with row vectors $\mathbf{x_i}$ for $i = 1,...,n$, PCA uses a set of k-dimensional weight vectors $\mathbf{w_f} = (w_1,...,w_k)$ to produce modified row vectors $\mathbf{y_i} = (y_1,...y_g)$ where $y_f = \mathbf{w_f} \cdot \mathbf{x_i}$ for $i = 1,...,n$ and $f = 1,...,g$. The weight vectors are defined such that the maximal variance of $\mathbf{X}$ is explained while each component y_f is orthogonal to all previous components.

This basic framework is used to determine the individual components and associated variable weights for both the household and municipality levels. Given a set of $k$household-level variables for N households distributed in m areal units, two data matrices are defined: a N-by-k household data matrix, and a m-by-$\tilde{k}$ aggregate data matrix where $\tilde{k}$ are the municipality-upscaled equivalents of k. The upscaled versions are calculated as regional means for each variable. while it might seem counterintuitive to upscale household-level data as opposed to simply using other regional data sources, doing so ensures that both levels are more directly comparable.

Performing PCAs on each these data matrices produces unique components and weighting vectors $\mathbf{w_{f_h}}$ and $\mathbf{w_{f_m}}$ for households and municipalities, respectively. In order to maximize comparability with previous vulnerability literature, component loadings are calculated with a Varimax orthogonal rotation and Kaiser normalization. The Varimax rotation improves the interpretability of components by maintaining orthogonality while ensuring that variables are strongly loaded on a single component and as close to zero as possible on the others. Subsequently, variable loadings are used to contextually interpret each of the components. These component results are used to determine whether first-order qualitative or quantitative differences between scale levels exist. Some differences are to be expected on the basis of variable correlations increasing with aggregation level, but large qualitative differences would point to a lack of measurement equivalence between scales. To further facilitate comparison between levels, the correlations between regionally averaged household component scores are also calculated. These correlations indicate whether uncorrelated household dimensions become interrelated at larger scales (Härnqvist, 1978, Subramanian et al., 2009, Puntscher et al., 2016).

2.5.2 Regional composite index consistency

In most inductive indices, some number of principal components are combined to form composite vulnerability scores. In formal terms, the vulnerability index score for some units $i = 1,...,m$ is defined as a weighted sum of the components above some cutoff threshold:

$$SV_i = \sum_{f=1}^{g*} s_f \cdot y_{fi} \tag{2.1}$$

where $g*$ is the number of retained components and s_f is a vector of component weights. Usually dimensions are equally weighted, but some studies have used the percentage of total variance explained as a weighting factor instead.

This approach is used to calculate baseline social vulnerability index rankings for each municipality. Then, an alternative method is implemented where composite scores are first calculated at the household level (i.e. with weights $\mathbf{w_{f_h}}$) and then averaged to regions. The conceptual idea here is to—quite literally—define a given region's social vulnerability as the average vulnerability of all the households it contains. This is distinct from Eq. (1) in that index components are uncorrelated between households because the aggregation occurs after the PCA implementation. As a result, referring to the 'average household' in this case is not an instance of the ecological fallacy. Formally, for some units $i = 1,...,m$ each with n households, social vulnerability index scores are given by:

$$SV_i \sim \frac{\sum_{j=1}^{n} \sum_{f=1}^{g*} s_f \cdot y_{fj}}{n} \tag{2.2}$$

Both sets of social vulnerability scores are z-score standardized and each municipality assigned a rank based on its relative position within each index. These ranks are then compared with particular focus placed on identifying cases where ranks change significantly or index scores change sign. Like before, it is expected that there will be some variation in the scores produced with the two methods—one approach is maximizing variance across municipalities and the other across households. However, the degree to which this is the case is currently unknown. Evaluating rank changes provides a mechanism to characterize this particular scale sensitivity, informing further discussion over the relevance of aggregation issues for vulnerability index interpretation.

2.5.3 Household composite index consistency

A similar rank-change approach is implemented at the household level, leveraging equally sized weighting vectors. Baseline index scores are first produced for each household $(i = 1, ..., N)$ using Eq. (1) where $y_f = \mathbf{w_{f_h}} \cdot \mathbf{x_i}$. These are compared to alternative scores also calculated using Eq. (1), but the regional weighting vectors $\mathbf{w_{f_m}}$ are used such that $y_f = \mathbf{w_{f_m}} \cdot \mathbf{x_i}$. The standardized scores and ranks for each household are then compared both within municipalities and across the entire study area. The rationale for swapping weight vectors is to estimate household vulnerability scores along the dimensions of maximal variance at the regional level. This is conceptually similar to the approach taken in the previous section where composite household scores are averaged within regions.

While it is unlikely that this approach would be preferred for an actual household-level vulnerability assessment, it makes for a useful comparison. For example, it is plausible that maximizing variance between regions produces strong conceptual factors that do not arise from the household analyses but nonetheless are relevant at the household level. In this case, the differences between composite household scores reflect a household's relative ranking along different dimensions. Alternatively, it is possible that the regional components are simply noisy versions of the household components. In this situation, the rank differences between composite household scores are more akin to error measurements. Either way, comparing rank changes at the household level provides another way to evaluate differences between aggregated and disaggregated versions of the same data.

2.6 Results

2.6.1 Principal component analysis results

The PCA results at the household level are shown in Table 2.1. Six components are identified using an eigenvalue threshold of 1.0, capturing 76% of the total variance in the data spread roughly evenly across all six components. Each of the components has relatively straightforward interpretations derived from the loading variables, thematically consistent with previous studies on vulnerability-related themes in Nepal (Gentle and Maraseni, 2012, Muzzini and Aparicio, 2013, Tamang et al., 2014, Dewan, 2015, Jaquet et al., 2016, Mainali and Pricope, 2017). The components are described as follows:

1. **Old Age:** Loaded positively by households with elderly household heads or other elderly family members present in household.

2. **Migration:** Loaded positively by households with fully present members (lack of internal migration) and negatively by households with a member living abroad. Affected positively to a lesser degree by large households.

3. **Income-Education:** Loaded positively by households with access to a bank account, high household incomes, and educated household heads.

4. **Children:** Loaded positively by households with children (<5 years of age) present. Affected positively to a lesser degree by large households.

5. **Female Head of Household (HOH):** Loaded extremely strongly by female household heads. Weakly loaded by large households and household members living abroad.

6. **Disability:** Loaded almost exclusively by households with one or more disabled member(s) present.

Table 2.1: Varimax rotated component matrix for PCA of household data.

n=741,067	Component					
	1	2	3	4	5	6
Gender	0.03	-0.07	-0.04	0.01	**0.92**	0.00
Income Level	-0.02	-0.34	**0.67**	-0.09	0.10	0.06
Education Level	0.47	0.13	**0.65**	-0.02	-0.23	0.03
Size of Household	0.13	**0.59**	0.03	**0.53**	0.32	0.09
Bank Account	-0.02	0.04	**0.81**	0.08	-0.02	-0.03
Household Head Age	**0.84**	0.10	0.17	-0.20	0.10	0.02
Members Abroad	0.02	**-0.70**	0.05	-0.11	0.33	0.02
Household Presence	-0.01	**0.83**	-0.07	-0.24	0.07	-0.00
Disabilities	0.04	0.01	0.02	0.00	0.01	**1.00**
Children Presence	-0.14	-0.09	0.00	**0.91**	-0.04	-0.02
Elderly Presence	**0.85**	-0.09	-0.07	0.05	-0.01	0.02
Variance Explained (%)	15	15	14	11	10	9
Cumulative Variance Explained (%)	15	31	45	56	67	76

Bolded values values > 0.5

The Pearson's correlation coefficients between the household components aggregated to Nepal's municipality and district levels are shown in Figure 2.2. These associations highlight the specific dimensions along which household components become correlated at aggregated levels. Although the effect sizes are difficult to compare across scale levels, some general trends emerge. At the municipality level, regional correlations are most evident for the income-education and old age dimensions. Income-education becomes positively correlated with the children and female

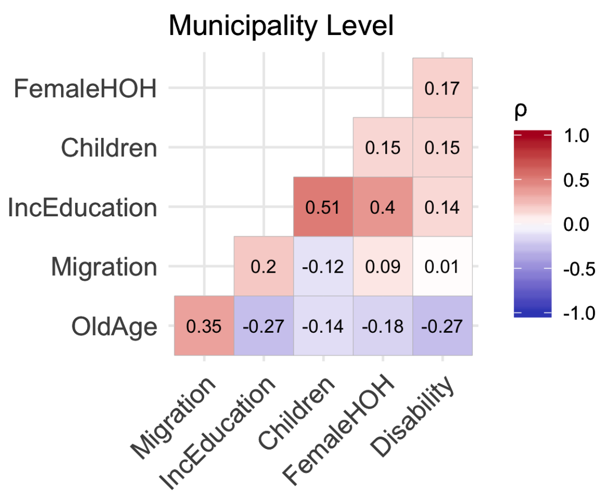

Figure 2.2: Pearson's correlation coefficient of household component scores averaged to municipality and district boundaries.

head of household dimensions while old age becomes moderately correlated with migration and income dimensions. At the district level, almost all of the dimensions become correlated to some degree, suggesting that household components (with the exception of disability status) would be difficult to extract from district-level analyses.

The PCA results at the municipality level are shown in Table 2.2, with six components explaining 89% of the variance in the data. Considering the strongest variable loadings, the components are in large part qualitatively similar to those at the household level (Table 2.2). However, there are quite a few smaller changes in variable loadings, explained by the component correlations in Figure 2.2. For example, household presence transitions from strongly loading the migration dimension in the household results to weakly loading several dimensions in the municipality results. Similarly, presence of children has zero contribution to the income-education dimension at the household level compared to a 0.3 loading at the municipality level. It is not clear, given the extent of these small differences, whether the municipality and household components should be considered equivalent.

Table 2.2: Varimax rotated component matrix for PCA of municipality data.

n=110	Component					
	1	2	3	4	5	6
Gender	-0.13	0.10	0.08	0.20	**0.92**	0.09
Income Level	-0.04	**0.74**	0.27	0.26	0.01	0.36
Education Level	-0.09	**0.62**	0.40	-0.36	0.31	-0.06
Size of Household	0.18	0.27	**0.84**	-0.08	0.16	0.08
Bank Account	0.06	**0.88**	0.14	-0.05	0.02	-0.02
Household Head Age	**0.95**	0.00	-0.16	-0.09	-0.07	-0.07
Members Abroad	-0.18	-0.02	-0.07	**0.90**	0.27	-0.07
Household Presence	**0.60**	-0.04	-0.36	-0.49	0.34	0.20
Disabilities	-0.10	0.09	0.08	-0.09	0.08	**0.95**
Children Presence	-0.46	0.30	**0.80**	0.08	-0.11	0.11
Elderly Presence	**0.90**	0.03	0.17	-0.09	-0.15	-0.09
Variance Explained (%)	22	17	16	12	11	10
Cumulative Variance Explained (%)	22	39	55	67	78	89

Bolded values values $>$ 0.5

2.6.2 Municipality composite score rank changes

As mentioned previously, this paper is focused on comparing the methodological approaches and is not intended to be a vulnerability assessment. Nevertheless, index results are mapped to assist in visualizing and interpreting results. The standardized composite social vulnerability index scores calculated at the municipality level using both methodologies described in section 4.2 are shown in Figure 2.3. Positive scores (red) indicate higher degrees of social vulnerability, relative to the study area and variables included in the analysis.

There is some similarity between the two approaches' results in terms of broad spatial trends of positive and negative scores. In general, southern and eastern municipalities tend to score as more vulnerable than those in the north and west. In terms of specific municipality scores

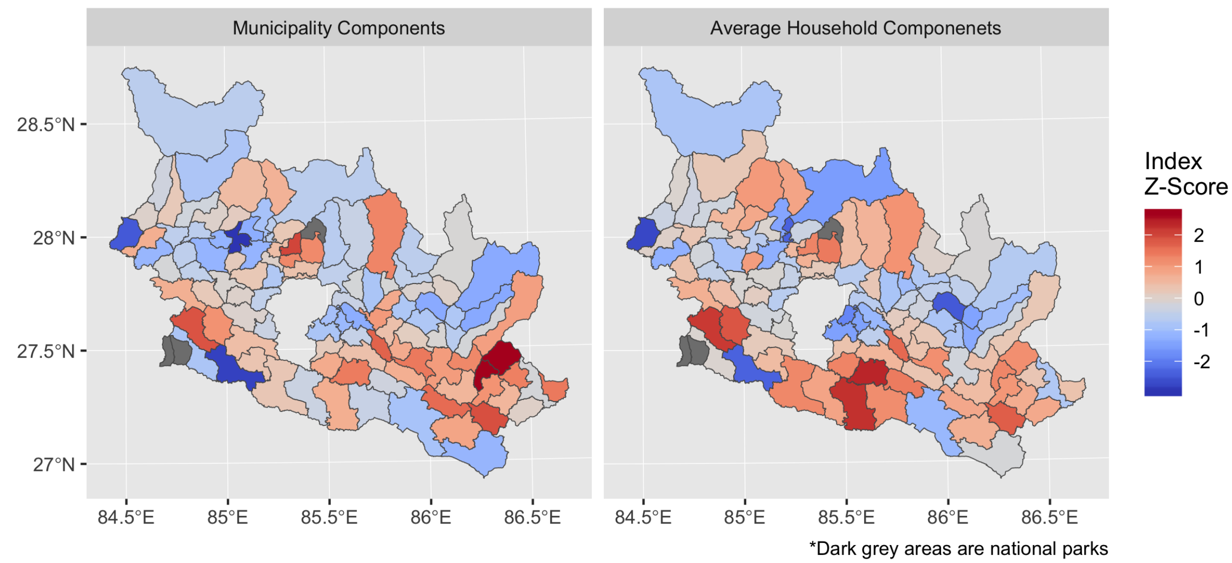

Figure 2.3: Composite social vulnerability index scores for each municipality calculated with equally weighted municipality PCA component scores (left) and municipality averaged household composite scores (right).

however, there are notable differences between the two approaches. These differences are quantified in Figure 2.4, with specific focus placed on magnitude and sign changes. Here, a magnitude classification is assigned to a municipality with a relative rank change of greater than 20% (22 positions). A sign classification identifies cases where municipalities switch from positive index scores to negative index scores or vice versa. These classifications provide a baseline set of criteria by which to identify cases where index interpretations might plausibly vary between scores.

Of the 110 municipalities in the study area, 37% are classified with either magnitude and sign changes, or both. If the magnitude percent threshold is relaxed to 10% or increased to 30%, the percentages change to 57% and 22%, respectively. Accordingly, for anywhere between one fifth and one half of municipalities, index rankings generated from municipality aggregated data are substantially different from their household equivalents. Thus, while the scores from both methods are correlated, the relative ranking of municipalities is sensitive to the aggregation scale

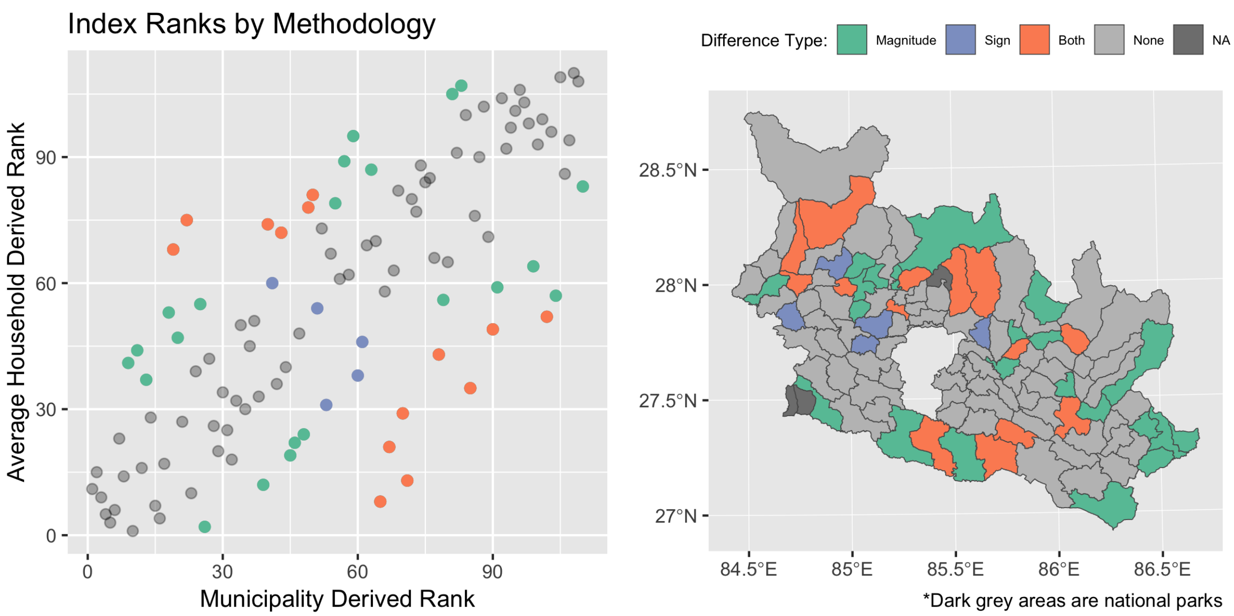

Figure 2.4: Classification of potential interpretation issues due to rank differences in magnitude, sign, or both. Municipalities plotted by rank (left) and mapped by classification type (right).

of the data. In context with Tables 2.1 and 2.2, these results also provide evidence that even qualitatively similar components can produce quite different composite index scores.

2.6.3 Household composite score rank changes

The distribution of standardized household-level composite scores calculated according to section 4.3 are shown in Figure 2.5. The composite scores are grouped by municipality and visualized as overlain density curves due to the size of the household data set. Each curve shows the relative distribution of household composite scores in a given municipality. The distributions using the household weighting vectors are slightly positively skewed with a median just below zero (-0.15). This particular shape makes sense in the context of vulnerability where the 'most vulnerable' households occur in low probabilities but are significantly more vulnerable than other households on the spectrum. When using the municipality weighting vectors, the composite scores are bi-modally normally distributed with symmetric peaks just above and below zero. In both cases the distribution shapes are similar across municipalities, indicating that the demographic distributions between municipalities are alike.

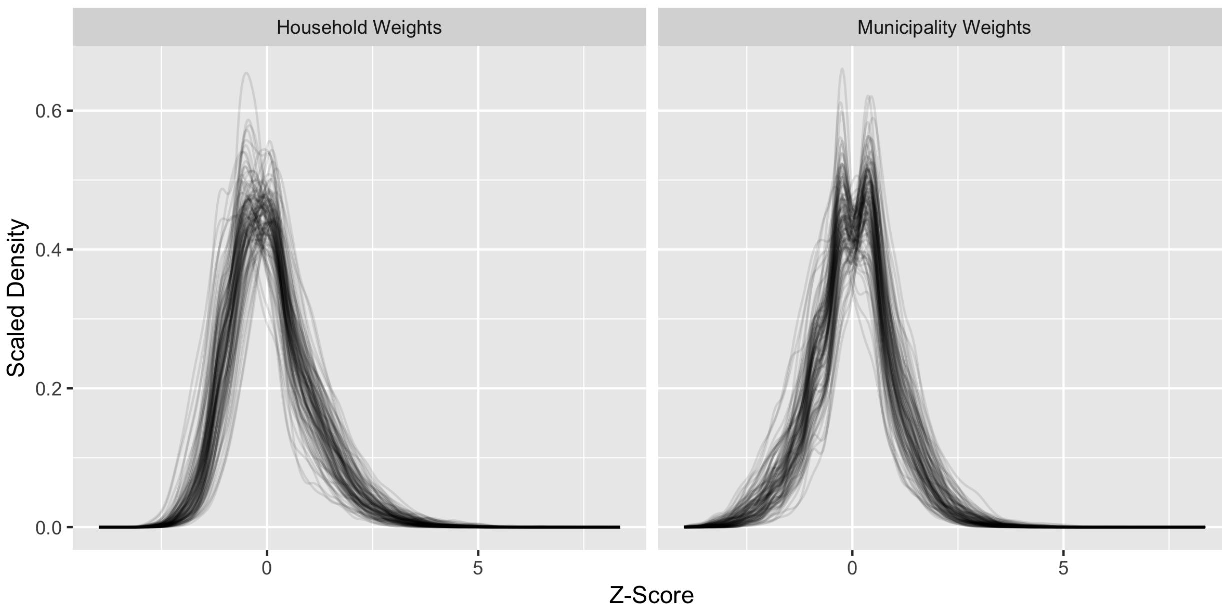

Figure 2.5: Density curves of household level composite social vulnerability index scores, grouped by municipality.

The differences between scores produced with the two different weighting vectors are seen in Figure 2.6 as both z-score differences and index rank changes for each household relative to its municipality. The percent rank change specifies where a household would be positioned on the index relative to all the other households in the same municipality. The median z-score change is 0.16, with a maximum positive and negative changes of 1.99 and -3.43, respectively. In terms of ranks, the median change is -3%, with maximum positive and negative changes of 90% and -76%, respectively. The rank change distributions are heavy-tailed, indicating that while most households are not significantly affected by the weighting adjustment, those that are affected are impacted to a great degree. For example, the maximum rank changes imply a household could theoretically be characterized as the least vulnerable household along one weighting scheme and in the top 10% vulnerable households of another—even when both weighting schemes are derived from same variables. To an even greater degree than before, these results reflect significant index score sensitivities to changes in component variable loadings.

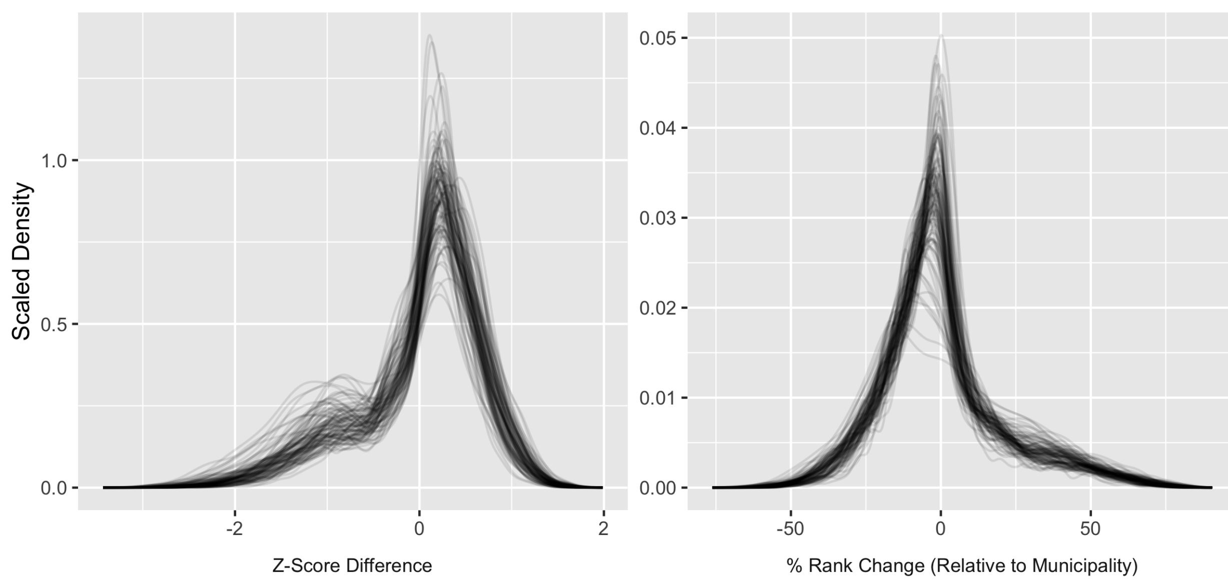

Figure 2.6: Density curves of differences in z-score (left) and rank (right) between household and municipality weighted composite scores.

The household rank changes are also visualized by mapping the percentage of households in each municipality classified with magnitude or sign changes according to the criteria introduced above. These results are shown in Figure 2.7. On the whole, anywhere between 21%-48% of households are classified with potential interpretation issues in any given municipality. More often than not, these issues are related to both sign and magnitude. Averaging across all municipalities, 6% of households have sign reversals, 12% of households see a rank change of 20% or more, and 15% experience both. There is some spatial variation in these percentages across the study area, but in general the results are consistent. These findings follow from the consistency of

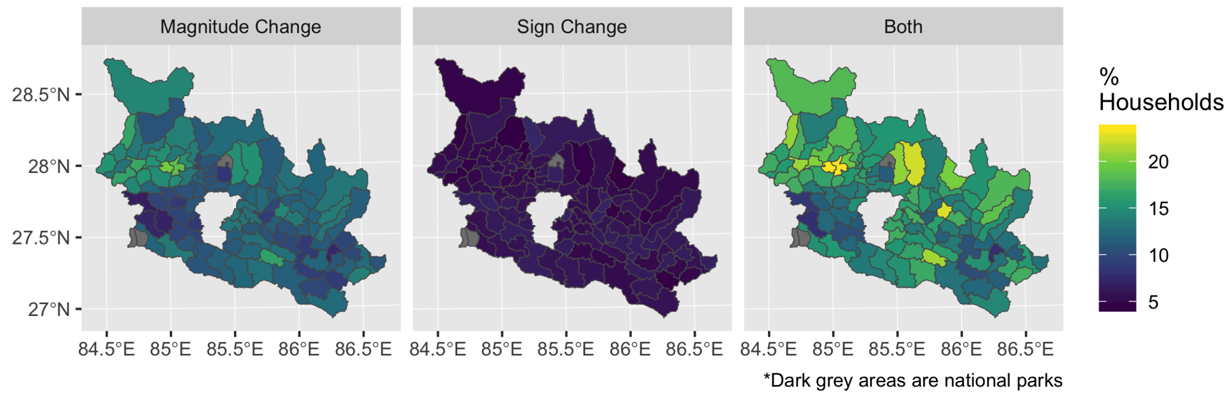

Figure 2.7: Percent of households per municipality that are classified as having potential interpretation errors in magnitude, sign, or both.

the household index distributions between municipalities (Figure 2.5).

2.7 Discussion & Conclusions

This study provides empirical results that help characterize the relevance of aggregation issues to modeling social vulnerability with inductive index methods. Section 5.1 showed that otherwise identically designed PCAs can produce different results at the household level (Table 2.1) and regional levels (Table 2.2). For the two scales considered in this study, these differences manifest as subtle quantitative differences in variable loadings rather than large qualitative changes. Similar trends were found in the PCA-based sensitivity analysis of the SoVI (Schmidtlein et al., 2008). Somewhat intuitively, the differences between components derived from household and regional levels reflect the correlations between household components when aggregated to regional levels. While these results are not particularly interesting by themselves, Sections 5.2 and 5.3 show that even small quantitative differences in component loadings can massively affect how regions or households are ranked with composite index scores. This suggests that qualitative similarity between vulnerability dimensions is not sufficient evidence to assume that subsequent index scores are similar.

It is not completely clear how these differences should be interpreted in the context of a vulnerability assessment, or at what magnitude they become relevant. Many previous vulnerability index studies constrain their interpretations to the scale at which the analysis was performed, assuming that index components reflect relevant processes operating at the given scale (Beccari, 2016). However, based on the results of this study, it is not clear whether the components derived from regional means indicate regional vulnerability dimensions or are simply coarse approximations of household characteristics. Figure 2.4 shows that a non-trivial number of municipalities rank quite differently if social vulnerabilities are calculated by maximizing variance along individual households rather than municipality averages. The same is true for households when the regional PCA weights are used instead of household weights (Figure 2.7). These are important results because they challenge the practice of using local-scale studies to justify the associations of variables at aggregated scales. As soon as data is stratified into regions (or any type of subpopulation), there is no guarantee that the associations between variables maintain their directions and magnitudes.

That is not to say that a pure household-level analysis should be the preferred approach. The notion of scale hierarchies is foundational to vulnerability theory (Turner et al., 2003, Wisner et al., 2004, Birkmann, 2007, Fekete et al., 2010). Analyzing a pooled version of household data potentially neglects the regional characteristics that form the back bone of many vulnerability studies, particularly those using a ‘vulnerability of places’ based approach. Thus, the question here is not whether regional characteristics should be considered, but whether aggregated data on household characteristics are sufficiently strong indicators for such factors. Framed another way, it is hard to determine at what point a particular series of variables stops measuring one construct and starts measuring another. For example, in a household-level case study in Western Nepal, Thieme and Wyss (2005) found that labor migration contributes to livelihood sustainability and increases financial capital, educational attainment of children, and social capital. On the other hand, Sunam and McCarthy (2016) argues that labor migration in Nepal at large contributes to structural processes that deepen inequality. Although these types of differences could technically be identified with the present methods, it is not obvious that aggregating household data produces clear indicators of regional processes. Household-level index components are correlated at regional levels to some degree (Figure 2.2), but it is difficult to evaluate whether these correlations are meaningful, spurious, or some combination of both.

In Nepal and other countries with abundant social and development programs, these findings are particularly relevant for government agencies and NGOs looking to incorporate measures of vulnerability into their decision making frameworks. Based on the results of this study, it is not clear that targeting an intervention at a highly socially vulnerable municipality would include the most socially vulnerable households ranked along the same criteria. While this variability may or may not be important in different contexts, it should be understood that socio-demographic variability within municipal or other administrative units can affect the trends seen in aggregated indices. These reverse concept—generalizing local household vulnerabilities to larger regions—is also important with respect to pilot programs and other local initiatives. Choosing variables for a regional social vulnerability assessment based on select household outcomes alone may not be sufficient for characterizing broader processes. As a general rule, quantitative assessments of vulnerabilities would benefit from more attention being placed on evaluating the scale consistency of proxy variables against the constructs they intend to represent. This is especially true in places like Nepal that are undergoing rapid social changes dominated by multi-scalar processes like labor migration and urbanization. Other studies have come to similar conclusions, recommending the use of qualitative, local-scale case studies as a means to validate sub-national vulnerability indices and ensure that results are consistent with local contexts (Schmidtlein et al., 2008, Fekete, 2009, Fekete et al., 2010).

The alignment of vulnerability theory and empirical methods has been an relevant subject of discussion in DRR literature for some time now (Cutter et al., 2000, Rashed and Weeks, 2003, Birkmann, 2007, Jones and Andrey, 2007, Barnett et al., 2008, Tate, 2013). Yet, as Fekete et al. (2010) notes, conceptual models of vulnerability are often vague with respect to how different scales should be handled. As a result, vulnerability indices have predominantly been tailored for application at a single scale despite underlying statistical concerns (Beccari, 2016). Using a micro-data set of over 740,000 households in rural Nepal, this study implemented a comparison of inductive index results at the household and municipality levels. On the whole, this analysis re-echoes existing concerns in the literature over variable contextualization. Even with a large micro-data set, determining which components adequately represent vulnerabilities and which result from aggregation effects is not a straightforward process. Small quantitative differences at this stage can result in large differences in composite scores, further complicating the interpretation process. These findings add to a growing body of literature examining issues of sensitivity and uncertainty in the measurement of vulnerability. While this paper considers aggregation issues in the context of an inductive index design, the results have implications for other index designs as well. Because deductive and hierarchical index structures involve similar decisions over analysis scale, both approaches also face potential issues of changing variable associations. Scale specification is also indirectly linked to variable selection and representation, issues that have received relatively more focus in the index literature (Jones and Andrey, 2007, Schmidtlein et al., 2008, Tate, 2012, 2013, Rufat et al., 2015). Choosing to include a particular variable does not necessarily mean the underlying construct is well represented. On the whole, this research furthers understanding on the representation of vulnerabilities in aggregated and disaggregated contexts, but there remain significant methodological challenges in the treatment of vulnerability as a multi-scalar process. Continued work in this area is likely to remain a promising future direction—especially as more micro-data sets become available in the future.

2.8 Acknowledgments

This research was performed under a National Science Foundation Graduate Research Fellowship, grant no. DGE-1450079. Three reviewers were especially helpful in improving the quality of the manuscript.

2.9 Works Cited

Anbarci, N., Escaleras, M., and Register, C. Earthquake fatalities: the interaction of nature and political economy. *Journal of Public Economics*, 89:1907–1933, 2005.

Barnett, J., Lambert, S., and Fry, I. The Hazards of Indicators: Insights from the Environmental Vulnerability Index. *Annals of the Association of American Geographers*, 98(1):102–119, Feb. 2008. ISSN 0004-5608. doi: 10.1080/00045600701734315.

Beccari, B. A Comparative Analysis of Disaster Risk, Vulnerability and Resilience Composite Indicators. *PLoS Currents*, 2016. ISSN 2157-3999. doi: 10.1371/currents.dis.453df025e34b682e9737f95070f9b970.

Birkmann, J. Risk and vulnerability indicators at different scales: Applicability, usefulness and policy implications. *Environmental Hazards*, 7(1):20–31, jan 2007. ISSN 1747-7891. doi: 10.1016/J.ENVHAZ.2007.04.002.

Borden, K. A., Schmidtlein, M., Emrich, C., and Cutter, S. L. Vulnerability of U.S. Cities to environmental hazards. *Journal of Homeland Security and Emergency Managment*, 4(2): 229–248, 2007.

Burton, C. G. Social Vulnerability and Hurricane Impact Modeling. *Natural Hazards Review*, 11 (2):58–68, 2010. doi: 10.1061/(ASCE)1527-6988(2010)11:2(58).

Chakraborty, J., Tobin, G. A., and Montz, B. E. Population Evacuation: Assessing Spatial Variability in Geophysical Risk and Social Vulnerability to Natural Hazards. *Natural Hazards Review*, 6(1):23–33, 2005. doi: 10.1061/(ASCE)1527-6988(2005)6:1(23).

Clark, W. a. V. and Avery, K. L. The Effects of Data Aggregation in Statistical Analysis. *Geographical Analysis*, 8(4):428–438, 1976. ISSN 1538-4632. doi: 10.1111/j.1538-4632.1976.tb00549.x.

Cutter, S. L., Mitchell, J. T., and Scott, M. S. Revealing the vulnerability of people and places: A case study of Georgetown County, South Carolina. *Annals of the American Association of Geographers*, 90(4):713–737, 2000.

Cutter, S. L., Boruff, B. J., and Shirley, W. L. Social Vulnerability to Environmental Hazards. *Social Science Quarterly*, 84(2):242–261, jun 2003. ISSN 0038-4941. doi: 10.1111/1540-6237.8402002.

Damm, M. Mapping Social-Ecological Vulnerability to Flooding - A sub-national approach for Germany. page 226, 2010.

[dataset] Kathmandu Living Labs. 2015 Nepal Earthquake: Open Data Portal. URL https://v1-opendata.klldev.org/{\#}/.

Dewan, T. H. Societal impacts and vulnerability to floods in Bangladesh and Nepal. *Weather and Climate Extremes*, 7:36–42, 2015. ISSN 2212-0947. doi: 10.1016/j.wace.2014.11.001.

DeYoung, S. E. and Penta, S. Issue Attention and Group Mobilization for Caste Rights Following the 2015 Gorkha, Nepal, Earthquake. *Earthquake Spectra*, 33(S1):S403–S414, Oct. 2017. ISSN 8755-2930. doi: 10.1193/120716EQS225M.

Fafchamps, M. and Shilpi, F. Determinants of the Choice of Migration Destination. Discussion Paper 7407, Centre for Economic Policy Research, London, 2009.

Fekete, A. Validation of a social vulnerability index in context to river-floods in Germany. *Natural Hazards and Earth System Sciences*, 9(2):393–403, Mar. 2009. ISSN 1684-9981. doi: 10.5194/nhess-9-393-2009.

Fekete, A., Damm, M., and Birkmann, J. Scales as a challenge for vulnerability assessment. *Natural Hazards*, 55(3):729–747, 2010. ISSN 0921030X. doi: 10.1007/s11069-009-9445-5.

Gartaula, H., Niehof, A., and Visser, L. Shifting perceptions of food security and land in the context of labour out-migration in rural Nepal. *Food Security*, 4:181–194, 2012.

Gentle, P. and Maraseni, T. N. Climate change, poverty and livelihoods: adaptation practices by rural mountain communities in Nepal. *Environmental Science & Policy*, 21(Supplement C): 24–34, nov 2012. ISSN 1462-9011. doi: 10.1016/j.envsci.2012.03.007.

Ghimire, S. Data Revolution, 2016 (accessed October 7, 2018). URL http://admin.myrepublica.com/the-week/story/43132/banking-on-data.html.

Härnqvist, K. Primary Mental Abilities at collective and individual levels. *Journal of Educational Psychology*, 70(5):706–716, 1978. doi: 10.1037/0022-0663.70.5.706.

He, L., Aitchison, J. C., Hussey, K., Wei, Y., and Lo, A. Accumulation of vulnerabilities in the aftermath of the 2015 Nepal earthquake: Household displacement, livelihood changes and recovery challenges. *International Journal of Disaster Risk Reduction*, 31:68–75, Oct. 2018. ISSN 22124209. doi: 10.1016/j.ijdrr.2018.04.017.

Jaquet, S., Shrestha, G., Kohler, T., and Schwilch, G. The effects of migration on livelihoods, land management, and vulnerability to natural disasters in the Harpan Watershed in Western Nepal. *Mountain Research and Development*, 36(4):494–505, 2016.

Jones, B. and Andrey, J. Vulnerability index construction: methodological choices and their influence on identifying vulnerable neighbourhoods. *International Journal of Emergency Management*, 4(2):269, 2007. ISSN 1471-4825. doi: 10.1504/IJEM.2007.013994.

Keefer, P., Neumayer, E., and Plümper, T. Earthquake propensity and the politics of mortality prevention. *World Development*, 39(9):1530–1541, 2011.

Khazai, B., Kunz-Plapp, T., Büscher, C., and Wegner, A. VuWiki: An Ontology-Based Semantic Wiki for Vulnerability Assessments. *International Journal of Disaster Risk Science*, 5(1): 55–73, mar 2014. ISSN 2095-0055. doi: 10.1007/s13753-014-0010-9.

Mainali, J. and Pricope, N. G. High-resolution spatial assessment of population vulnerability to climate change in Nepal. *Applied Geography*, 82:66–82, May 2017. ISSN 01436228. doi: 10.1016/j.apgeog.2017.03.008.

Muzzini, E. and Aparicio, G. *Urban growth and spatial transition in Nepal: An initial assessment*. The World Bank, Washington, D.C., 2013.

Neumayer, E., Plümper, T., and Barthel, F. The political economy of natural disaster damage. *Global Environmental Change*, 24(1):8–19, 2014.

NSET. Gorkha Earthquake: Loss Data and Figures, 2015. URL http://www.nset.org.np/eq2015/.

Openshaw, S. *The Modifiable Areal Unit Problem*. Geo Books, 1983. ISBN 978-0-86094-134-7. Google-Books-ID: BJ5jQgAACAAJ.

Paudel, K. P., Tamang, S., and Shrestha, K. K. Transforming land and livelihood: Analysis of agricultural land abandonment in the mid hills of Nepal. *Journal of Forest and Livelihood*, 12 (1):11–19, 2014.

Phillips, B. D., Thomas, D. S. K., Lovekamp, W. E., Fothergill, A., Phillips, B. D., Lovekamp, W. E., and Fothergill, A. *Social Vulnerability to Disasters*. CRC Press, May 2013. ISBN 978-1-4665-1638-0. doi: 10.1201/b14854.

Puntscher, S., Hauser, C., Walde, J., and Tappeiner, G. Measuring Social Capital with Aggregated Indicators: A Case of Ecological Fallacy? *Social Indicators Research*, 125(2):431–449, jan 2016. ISSN 0303-8300. doi: 10.1007/s11205-014-0843-z.

Rashed, T. and Weeks, J. Assessing vulnerability to earthquake hazards through spatial multicriteria analysis of urban areas. *International Journal of Geographical Information Science*, 17(6):547–576, 2003.

Robinson, W. S. Ecological Correlations and the Behavior of Individuals. *American Sociological Review*, 15(3):351, jun 1950. ISSN 00031224. doi: 10.2307/2087176.

Rufat, S., Tate, E., Burton, C. G., and Maroof, A. S. Social vulnerability to floods: Review of case studies and implications for measurement. *International Journal of Disaster Risk Reduction*, 14:470–486, dec 2015. ISSN 2212-4209. doi: 10.1016/J.IJDRR.2015.09.013.

Rygel, L., O'Sullivan, D., and Yarnal, B. A method for constructing a social vulnerability index: An application to hurricane storm surges in a developed country. *Mitigation and Adaption Strategies for Global Change*, 11:741–764, 2006.

Schmidtlein, M. C., Deutsch, R. C., Piegorsch, W. W., and Cutter, S. L. A sensitivity analysis of the social vulnerability index. *Risk Analysis*, 28(4):1099–1114, 2008. ISSN 02724332. doi: 10.1111/j.1539-6924.2008.01072.x.

Schmidtlein, M. C., Shafer, J. M., Berry, M., and Cutter, S. L. Modeled earthquake losses and social vulnerability in Charleston, South Carolina. *Applied Geography*, 31(1):269–281, Jan. 2011. ISSN 0143-6228. doi: 10.1016/j.apgeog.2010.06.001.

Shapira, S., Aharonson-Daniel, L., and Bar-Dayan, Y. Anticipated behavioral response patterns to an earthquake: The role of personal and household characteristics, risk perception, previous experience and preparedness. *International Journal of Disaster Risk Reduction*, 31:1–8, Oct. 2018. ISSN 2212-4209. doi: 10.1016/j.ijdrr.2018.04.001.

Subramanian, S. V., Jones, K., Kaddour, A., and Krieger, N. Revisiting Robinson: The perils of individualistic and ecologic fallacy. *International Journal of Epidemiology*, 38(2):342–360, apr 2009. ISSN 1464-3685. doi: 10.1093/ije/dyn359.

Sunam, R. K. and McCarthy, J. F. Reconsidering the links between poverty, internal labour migration, and agrarian change: critical insights from Nepal. *The Journal of Peasant Studies*, 43(1):39–63, 2016.

Tamang, S., Paudel, K. P., and Shrestha, K. K. Feminization of Agriculture and its implications for food security in rural Nepal. *Journal of Forest and Livelihood*, 12(1):20–32, 2014.

Tate, E. Social vulnerability indices: a comparative assessment using uncertainty and sensitivity analysis. *Natural Hazards*, 63:325–347, 2012. doi: 10.1007/s11069-012-0152-2.

Tate, E. Uncertainty Analysis for a Social Vulnerability Index. *Annals of the Association of American Geographers*, 103(3):526–543, may 2013. ISSN 0004-5608. doi: 10.1080/00045608.2012.700616.

Thieme, S. and Wyss, S. Migration patterns and remittance transfer in Nepal: a case study of Sainik Basti in Western Nepal. *International Migration*, 53(5):60–98, 2005.

Turner, B. L., Kasperson, R. E., Matson, P. A., McCarthy, J. J., Corell, R. W., Christensen, L., Eckley, N., Kasperson, J. X., Luers, A., Martello, M. L., Polsky, C., Pulsipher, A., and Schiller, A. A framework for vulnerability analysis in sustainability science. *Proceedings of the National Academy of Sciences*, 100(14):8074–8079, 2003.

Wisner, B., Blaikie, P., Cannon, T., and Davis, I. *At Risk: Natural hazards, people's vulnerability and disasters*. Routeldge, 2nd edition, 2004.

Wu, S.-Y., Yarnal, B., and Fisher, A. Vulnerability of coastal communities to sea-level rise: a case study of Cape May County, New Jersey, USA. *Climate Research*, 22:255–270, 2002.

2.10 Appendix

Table 2.3: Variable categories used in PCA computations.

Variable Name	**Scale-Level**	**Label(s)**
Gender of Household Head	1	Male
	2	Female
Average Monthly Income**	1	Rs 50,000 or more
	2	Rs 30,000-50,000
	3	Rs 20,000-30,000
	4	Rs 10,000-20,000
	5	Rs 10,000 or less
Education Level (Household Head)	1	Secondary
	2	Primary
	3	Pre-Primary
	4	Non-Formal
	5	Illiterate
Size of Household*	Continuous	NA
Bank Account Status	1	Has Bank Account
	2	No Bank Account
Age of Household Head	Continuous	NA
Household Members Abroad*	1	None
	2	>1 Member
Household Presence*	Percentage	
Disability Status*	1	No Disabled Members
	2	>1 Members
Children Presence*	1	No Children (<5 years old) Present
	2	Children (<5 years old) Present
Elderly Presence*	1	No Elderly (>65 years old) Present
	2	Children (<5 years old) Present

*Merged from individual table to household level
**Using conversion rate at time of submission, Rs 10,000 is equivalent to $86 USD

The eleven socio-demographic variables used for this analysis are shown in Table 3.2. R scripts for formatting the raw HRHRP data tables are available in this paper's corresponding code repository.

3: A BAYESIAN MODELING APPROACH FOR ESTIMATING EARTHQUAKE RECONSTRUCTION BEHAVIOR

This chapter corresponds to the following in-review paper: Wilson, B.S. (2019), A Bayesian modeling approach for estimating earthquake reconstruction behavior. *Annals of the American Association of Geographers*.

3.1 Abstract

Rebuilding and repairing damaged physical infrastructure is a primary source of disaster aid spending following major earthquakes. While aid distribution is monitored, it is not well understood how economic support and technical assistance affect reconstruction behavior. This study develops and evaluates a Bayesian Item Response Theory modeling framework for estimating the probability of reconstructive action from household-level survey data. Household responses on reconstruction status, aid received, and willingness to commit additional resources from Inter-Agency Common Feedback Project surveys (n=5913) collected 11, 12, and 14 months after the Gorkha, Nepal Earthquake are used to estimate the probability of reconstructive action. Results show differences in marginal reconstruction probabilities ranging from 2-78 percent across varying combinations of aid receipt and household willingness to commit additional resources. Estimated reconstruction probabilities are lowest for households with low willingness to commit additional resources and households that have not received a reconstruction-related engineering consultation. All model results showed strong variability with geographic location. These findings provide detailed quantitative estimates of earthquake recovery that have not previously been available and offer a promising methodology for utilizing future post-disaster household-

level survey data.

Bayesian modeling; item response theory; earthquake reconstruction; disaster recovery; Nepal Earthquake

3.2 Introduction

Effectively coordinating the inflow of aid resources is a major challenge for post-earthquake reconstruction efforts, especially in low and middle income countries where many households rely on disaster aid as their primary means for recovery. While aid distribution is increasingly monitored, the processes by which households undertake reconstruction actions are poorly understood. Previous research has shown that access to resources is a key determinant of housing reconstruction (Wu and Lindell, 2004, Epstein et al., 2018), but data from recent earthquakes shows large variations in reconstruction rates at the household level independent of disaster aid (Dunford and Li, 2011, Daly et al., 2017, He et al., 2018). Thus, it is important to better understand the connections between aid distribution and other household factors as they relate to reconstruction actions.

Current guidelines in the Sendai Framework for Disaster Risk Reduction recommend that reconstruction program standards, policies, and macro-level coordination are driven by a lead-recovery agency, while implementation of recovery activities take place by distributed actors at the local level (Bank, 2015). This quasi-decentralized approach promotes a 'local solutions to local problems' mindset where communities and households have ownership over the reconstruction progress. However, the extent to which local realities end up aligning with these goals is often debated, particularly with respect to material resource access and the availability of technical assistance—two necessary forms of support for many households (Daly et al., 2017, Hall et al., 2017, Mishra et al., 2017, Bownas and Bishokarma, 2018, He et al., 2018). Coordinating these support mechanisms is a major function of the lead-recovery agency, especially in the context of 'build back better' reconstruction programs that require adherence to seismic building codes. For this reason, household decisions to rebuild are influenced to some degree by both centralized and decentralized authority, even in owner-driven systems (Comfort and Joshi, 2017, Daly et al., 2017).

Household reconstruction progress is typically monitored in two primary ways. Aid disbursement is tracked and reported following the guidelines set forth in Post-Disaster Recovery Frameworks, covering the primary governmental support mechanisms (reconstruction grant funding, engineering consultations, material availability, etc.) (Bank, 2015). These statistics are useful for broadly evaluating the stage of reconstruction progress (Chang et al., 2010), but do not capture important heterogeneities at the household level—where reconstruction decisions are being made. These data are supplemented by household livelihood needs and perception surveys that assess whether local populations feel their reconstruction needs are being met. However, survey differences in demographics, aid access, and personal capabilities, among other factors, have limited comparative conclusions. Notably missing in the reconstruction space are any modeling approaches for transforming these types of data into actual estimates of reconstruction behavior.

This research uses data from the 2015 Gorkha, Nepal earthquake to develop and evaluate a Bayesian Item Response Theory (IRT) model framework for estimating earthquake reconstruction probabilities from survey data. For the most part, Nepal's reconstruction policies are consistent with the standards and recommendations described above. One of the primary components of Nepal's reconstruction framework is the Rural Housing Reconstruction and Recovery Program (HRRP), a multi-phase, multi-stakeholder project aimed at providing the technical and economic support necessary to guide the 'owner-driven' reconstruction process. A NPR 300,000 ($\sim$ \$3,000 USD) three tranche grant is disbursed to eligible households at different phases of rebuilding contingent upon adherence to specific construction standards. Beneficiaries are also eligible for technical and social support during the reconstruction progress in accordance with the HRRP standards. However, provision of this support is decentralized and only a small number of areas have received the full range of services. These differences in aid availability, along with variability in household characteristics, form the backbone of the model framework. Reconstruction ability is predicted at the household level using survey data on aid receipt and self-reported willingness to commit additional resources to reconstruction collected approximately one year after reconstruction activities began. Estimated ability levels subsequently interact with regionally variable item response functions to estimate the probability of reconstructive action.

3.3 The Gorkha Earthquake Reconstruction Framework

Nepals recovery from the Gorkha Earthquake has been slow and contentious. The first year of recovery was situated against a contentious political background that delayed reconstruction activities (Comfort and Joshi, 2017). The National Reconstruction Authority (NRA)the governing body designated to oversee the reconstruction process was originally established by political ordinance in late June of 2015, but failed to attain sufficient legal status and was suspended until formal legislation was passed in December 2015. The publication of the official Post Disaster Recovery Framework followed five months later in May of 2016. A lack of information on recovery policies, beneficiary eligibility, and new building codes, among other factors, limited reconstruction progress during this transitory period (Comfort and Joshi, 2017, Daly et al., 2017).

Within Nepals reconstruction framework, the NRA is responsible for allocating reconstruction aid to partner agencies and ministries, while project implementation is coordinated across several levels including 7 sub-regional offices, 14 district-level project implementation units, 160 local resource centers, and individual village development committee and ward offices (of Nepal, 2016). Local implementation units have significant authority in this system, including the ability able to partner with external donors and NGOs. Under NRA guidelines, most field-level reconstruction activitiesincluding grant disbursement and technical or social support-are coordinated by local institutions (Daly et al., 2017). Household eligibility for grant funding and other government support is determined from the results of official engineering damage assessments collected by Nepals Central Bureau of Statistics. The first tranche of grant funding is disbursed upon enrollment in the program, while the second and third tranches require that certain stages of reconstruction progress are verified as adhering to specific building designs or to the minimum standards of the national building code by an approved engineer.

Within these requirements, individual households are largely free to make their own reconstruction decisions, with local ward and village development committee offices serving as the primary access point for individual households seeking social or technical assistance during the reconstruction process. However, despite intentions to provide households with equal support through decentralized offices and local governments, the actual availability of reconstruction support has varied widely across the affected area. Local governments have the authority and onus to provide reconstruction support, but they often lack the capacity to effectively do so (Daly et al., 2017). Asymmetrical reconstruction governance, combined with shortages in engineers, has left many households without promised supporta key deficiency for those in areas that are resource poor and multi-dimensionally vulnerable (The Asia Foundation, 2017, Mishra et al., 2017, He et al., 2018).

3.4 Materials & Methods

3.4.1 Summary of reconstruction data

Household-level data on earthquake reconstruction and aid distribution were obtained from Nepal's Inter-Agency Common Feedback Project Surveys collected in April, May, and July of 2017. The Inter-Agency CFP, a joint effort organized by the United Nations Humanitarian Cluster for the Gorkha Earthquake, coordinated data collection and aggregation efforts across various partner agencies to provide a cohesive set of surveys in support of the entire humanitarian community. These surveys broadly address issues related to reconstruction perceptions, livelihood needs, food security, and protection concerns. These surveys were carried out in village development committees (henceforth VDCs—now municipalities under Nepal's new federal structure) across the 14 priority affected districts with 60% or greater of households eligible for governmental reconstruction funding. In each month, district level populations were used to proportionally allocate 2100 individual surveys to random VDCs using a probability proportionate to size methodology with a minimum of ten surveys collected per ward visited. A small percentage of surveys ($\sim$ 5%) for each iteration were saved to boost responses in low population districts. Survey enumerators were instructed to collect individual responses from a demographically diverse set of respondents to ensure a reflective sample. Additionally, twenty-five percent of survey responses were reserved for municipalities to ensure an adequate representation of urban versus rural populations. In total, the three survey months include responses for 5,913 households in 245 VDCs across the 14 districts.

The CFP surveys include a diverse set of information on the earthquake reconstruction process, including data on housing status, technical assistance and economic aid receipt, livelihood needs, and perceptions on varying aspects of the reconstruction process. For the purposes of this study, data on direct cash support, engineering consultations, and self-assessed ability to commit additional resources to reconstruction are used (details in Appendix A). These covariates capture most of the primary identified reasons for slow rebuilding rates one year into the reconstruction process (The Asia Foundation, 2017) while balancing computational complexity and minimizing the risk of overfitting. Figure 4.2 shows the percentages of households having taken reconstructive action across each level of aid distribution and household willingness to commit additional resources to reconstruction. The increase in reconstruction rates with both covariates is consistent with previous research emphasizing the importance of resource availability for household reconstruction progress (Wu and Lindell, 2004, Epstein et al., 2018). Case studies have suggested that in places like Nepal, households with alternative economic strategies (remittances, crop diversification, access to community-based resources, etc.) are more likely to rapidly progress through housing recovery (Chatterjee and Okazaki, 2018, Epstein et al., 2018).

3.4.2 Bayesian modeling framework

Originally developed for psychometric application in psychology and education research (Fox, 2010), IRT models describe the probability of a keyed response for a given individual to an item in question using an item response function that links a person's latent ability to item-specific parameters. Hierarchical extensions to IRT, including adopting random effects for groups or individuals, modeling item parameters as correlated (Glas and van der Linden, 2003), or adding predictor variables on ability (Alegana et al., 2017, 2018), allow for flexible model specifications (Fox, 2010, Sulis and Toland, 2017). Although IRT models are well established in other disciplines, they have not yet been applied in a disaster recovery context.

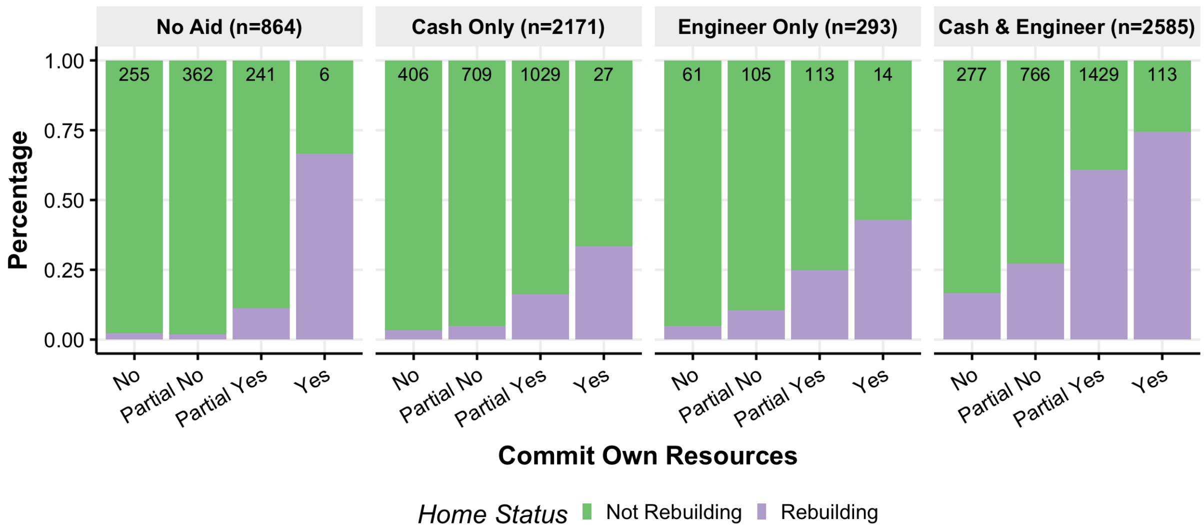

Figure 3.1: Reconstruction rates by aid level and household willingness to commit additional resources.

IRT models can be implemented in both frequentist and Bayesian frameworks, with both approaches seeing decades of application (Bock, 1972, Swaminathan and Gifford, 1985, 1986). This study adopts a Bayesian approach to more completely account for uncertainty in latent trait estimation and to clearly articulate the modeling assumptions. Reconstruction outcomes in an owner-driven system are expected to be somewhat uncertain and modeling full distributions are therefore valuable for interpreting results. Additionally, quantitatively modeling reconstruction is a new research direction and using a Bayesian framework allows for existing qualitative domain-knowledge to be incorporated into the model building process through the use of prior distributions (Gelman et al., 2017, Gabry et al., 2019). Defining a generative model is useful for incorporating specific contextual features of the reconstruction framework that are not present in the data itself.

Here, a unidimensional Bayesian IRT framework is used to estimate the probability of positive survey response for reconstruction actions, considering a single survey question on household-level housing status. For households $i = 1,...,I$ in districts $j = 1,...,J$ responding in survey month k, the probability of positive response ($Y_{ijk} = 1$) is modeled as:

$$P(Y_{ijk} = 1|\theta_i, a_{jk}, b_{jk}, c) = c + (1-c)\frac{exp[a_j(\theta_i - b_{jk})]}{[1 + exp[a_{jk}(\theta_i - b_{jk})]} \tag{3.1}$$

where Y_{ijk} is a dichotomous response variable reflecting whether a household i has started the reconstruction process, θ_i is the estimated latent ability parameter for household i, and a_{jk}, b_{jk}, c are the item discrimination, difficulty, and threshold parameters, respectively. In this context, the item discrimination parameter (a_{jk}) regulates the rate at which the probability of reconstruction initiation changes with different ability levels. The item difficulty parameter (b_{jk}) is equal to the point of median probability, describing the relative relationship between an item and the ability scale. The lower threshold parameter introduces a lower-bound on the response, allowing for households with low ability levels to have non-zero probability of positive response. For this application, variable intercepts are included on both the discrimination and difficulty parameters for both district and survey month. This accounts for the spatial and temporal variance in the probability of positive response at otherwise equal ability levels.

In a similar fashion to Fox and Glas (2001), Alegana et al. (2017, 2018), the ability parameter is modeled at a second level with a series of linear predictor variables ($X_1,...,X_Q$). Varying intercepts for VDCs ($v = 1,...,V$) are included to account for potential geographic effects:

$$
\begin{aligned}
&\theta_{iv} = \beta_0 + \beta_1 X_{1i} + \ldots + \beta_Q X_{Qi} + \alpha_v + e_{iv} \\
&e_{iv} \sim N(0, \sigma_e^2) \\
&\alpha_v \sim N(0, \sigma_\alpha^2).
\end{aligned}
\tag{3.2}
$$

with varying intercept α and regression coefficients β. The joint distribution on ability parameters is also multivariate normal with variance hyperparameters σ_e^2, and σ_α^2. Aid receipt (economic and technical) and self-reported ability to commit additional resources to reconstruction are included as predictor variables. Willingness to commit additional resources is modeled as an ordinal predictor variable while aid receipt is an unordered categorical predictor.

Proper prior distributions are specified for all parameters such that a prior marginal distribution exists for the data. To evaluate prior selection, a prior predictive simulation is implemented where MCMC draws are made from the joint prior distribution to evaluate whether simulated data is consistent with domain expertise. Following recommendations in (Gelman et al., 2017, Gabry et al., 2019), the objective here is to allow for improbable, but not impossible data to be drawn from the joint prior distribution Gabry et al. (2019). Several principles inform a realistic data generating process for the purposes of this study. First, the model is restricted such that item discrimination is positive. This helps with model identifiability, but also ensures that increases in reconstruction ability always result in increased reconstruction probability. Second, the lower threshold is specified to strongly favor values closer to zero based in an understanding that few rural households in Nepal have the financial or technical capacity to reconstruct their houses without assistance.

Weakly informative Student's T prior distributions with three degrees of freedom are used for the item discrimination and difficulty parameters with normal priors ($\sim N(0,1)$) on the standard deviations to constrain inter-district variability. As mentioned above, the prior on discrimination is truncated via an indicator function to constrain values to be positive. A beta distribution prior ($\alpha = 1, \beta = 10$) is placed on the lower-threshold parameter to constrain values between 0 and 1. The priors for coefficients on the θ parameter are treated separately to account for the difference in variable type. A weakly informative normal prior ($\sim N(0,1)$) is used for the aid receipt coefficients, and a Dirichlet prior is used for willingness to commit of additional resources coefficients with variable simplex parameters such that the change in response "partial no" to "partial yes" is a-priori more influential than changes from "no/partial no" or "yes/partial yes".

The Hamiltonian Monte Carlo (HMC) engine in Stan software (Carpenter et al., 2017), interfaced using the *brms* package in R (Brkner, 2017), is used to fit the model. HMC sampling has improved sampling efficiency for many hierarchical distributions compared to more traditional Markov-Chain Monte-Carlo (MCMC) methods like Gibbs sampling or Metropolis-Hastings. Convergence is evaluated with Stan's implementation of the potential-scale reduction factor ($\hat{R}$) and inspecting MCMC trace plots. Posterior predictive checks and leave-out-out cross-validation are used to evaluate model performance.

3.5 Results

The three parameter IRT model is fit using four MCMC chains running for 2000 iterations, 1000 of which are used as warm-up iterations. The model shows convergence on all parameters ($\hat{R}$ = 1.00) and does not show signs of any pathologies. Prior predictive simulation results, all parameter and uncertainty estimates, and model diagnostics are included in Appendix B. Validation was performed on a split 75-25 percent training-test data set to evaluate model calibration and accuracy (Appendix C). The data generated by the posterior predictive distribution (Figures 3.7, 3.8, 3.9) is consistently similar to the observed data and balanced across all predictor variables, districts, and survey months. In almost all cases, the observed data is centered within the 100 simulated predictions, indicating good model fit (Gabry et al., 2019). There are isolated cases (e.g. Kathmandu, July) where the distribution of simulated data does not contain the observed values, but on the whole the model appears well calibrated across grouping variables and does not show any clear biases towards particular combinations of covariate levels. These results suggest that the proposed model is able to generate accurate predictions for the probability of households starting the reconstruction process provided the requisite covariate data is available.

3.5.1 Reconstruction probabilities by ability level

Figures 2A and 2B shows the estimated median reconstruction probability along with corresponding variation by districts and 50%, 80%, and 95% quantile intervals across the ability spectrum. The median estimated ability levels for each combination of predictor variables are presented in 2C, with spatial variations by VDC variations shown in 2D. Estimated reconstruction probabilities are very low across the lower end of the ability range (lower-threshold 1%) corresponding to households that have not received an engineering consultation and do not have the ability to commit additional resources. Reconstruction probabilities rise over 50% at latent ability levels above two, generally corresponding to households that can fully commit additional resources (regardless of aid receipt) or households that can partially commit resources and have received both sources of aid. The uncertainties associated with reconstruction probabilities increase in the middle-range of the ability spectrum.

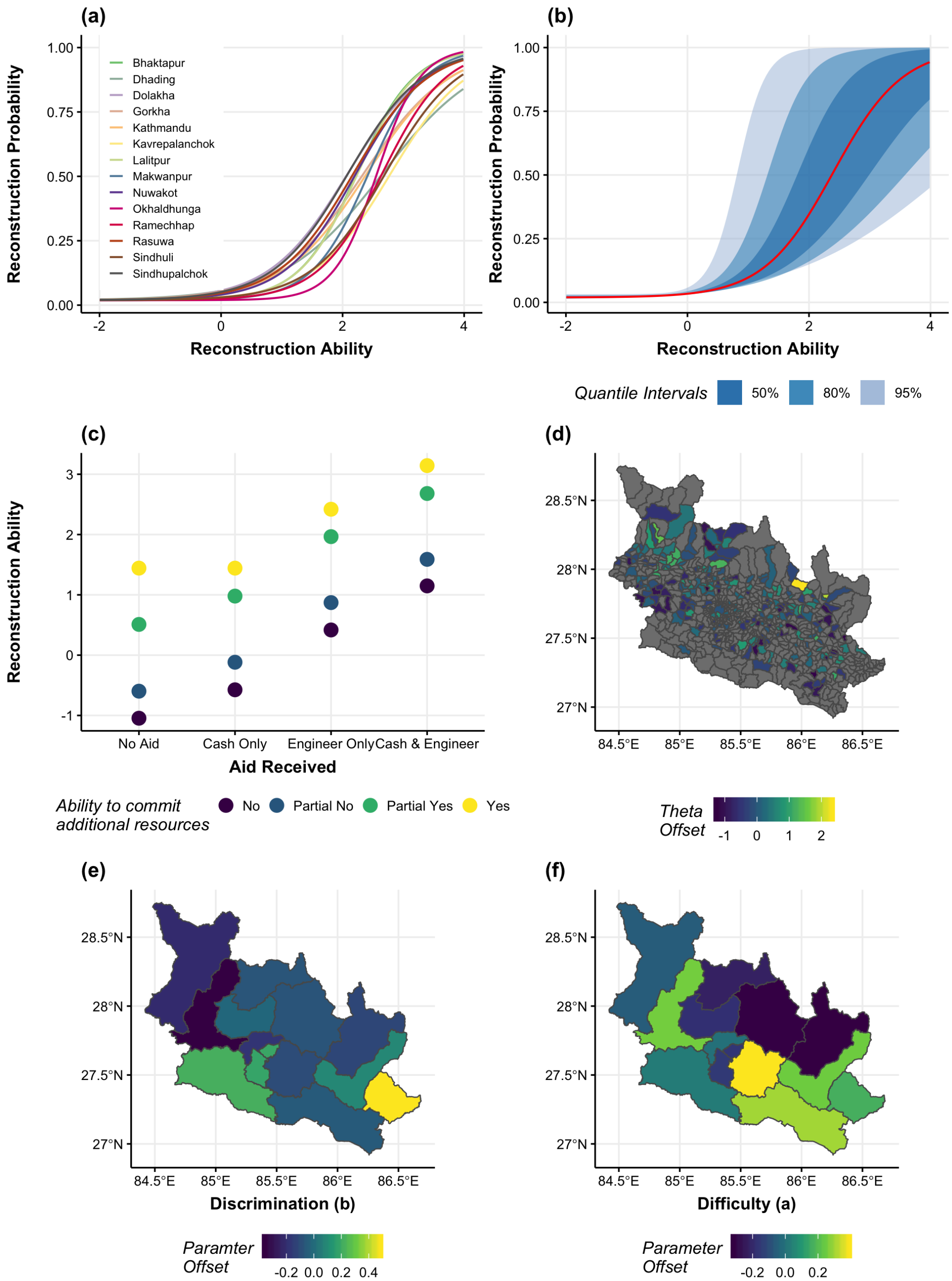

Figure 3.2: Posterior estimates for reconstruction probability. A: Item response curves by district, incorporating median random effects on difficulty and discrimination parameters. B: Population-level median item response curve with corresponding uncertainty windows. C: Median ability levels for each combination of predictor variables, marginalized across survey months. D: Median VDC level effects on estimated ability. E, F: Median District level variations in discrimination and difficulty parameters.

These trends are relative to a household's location. VDCs exert considerable influence on reconstruction ability, as shown in Figure 2D (standard deviation 0.79). Although some local clustering is present, the VDC variations do not indicate the presence of distance-decay or other spatial effects. Additionally, both discrimination and difficulty parameters vary by district (Figure 2E and 2F) with a variability of approximately 25% between minimum and maximum probabilities at equal ability levels. The highest and lowest difficulty parameter estimates are for Kavrepalanchok and Dolakha, respectively. Okhaldhuga and Dhading have the highest and lowest discrimination estimates.

3.5.2 Marginal effects of predictor variables

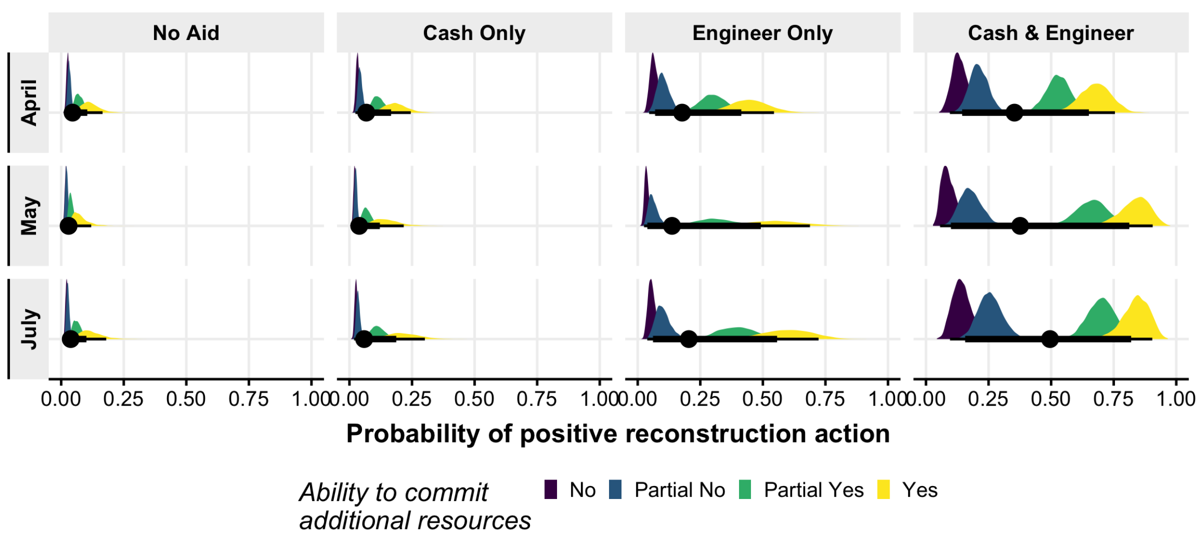

Figure 3.3: Estimated marginal effects of predictor variables. The black point and interval bars summaries the median posterior estimate and corresponding 50% and 95% uncertainty intervals. The colored distributions segment the densities by additional resource contribution.

Figure 3.3 shows the marginal effects on reconstruction probability for each of the predictor variable combinations across all three survey months. Although there are slight differences in the probabilities between the survey months, the same trends are seen across predictor variables. Comparing across aid receipt differences, the probability of positive reconstruction action without an engineering consultation is very low across all months and levels of additional household resource contribution. If households in this position are not able to commit additional resources, the estimated probability of reconstruction action(s) is effectively zero. These probabilities increase as households are able to commit additional resources, but the overall estimated probabilities remain low.

The additional resource dimension shows measurable gaps at every level of aid receipt, with the largest difference between partial no and partial yes responses. The effect of increasing additional resource contribution is quite strong, equivalent to or greater than receiving additional economic or technical aid. Notably, the resource gap increases as households receive more aid, suggesting that government assistance might have a greater benefit to those households already in an advantageous position. These differences evolve differently over time, with probabilities increasing for households able to commit additional resources while probabilities remain relatively stable for households unable to commit additional resources. These differences underscore the necessity of including additional information beyond aid receipt to characterize reconstruction progress.

3.6 Discussion & Conclusions

Monitoring household-level reconstruction progress is an important element of managing the rebuilding and recovery process in owner-driven frameworks. Yet, no previous studies have attempted to provide quantitative estimates of reconstructive action. Using household-level data from reconstruction livelihood and perception surveys collected after the 2015 Nepal Earthquake, this study finds considerable variation in the probability of households starting the rebuilding process along several dimensions, including aid receipt, self-reported willingness to commit additional resources, and geographic location. These results are largely consistent with qualitative findings from the social-science literature on earthquake reconstruction (Wu and Lindell, 2004, Chatterjee and Okazaki, 2018, Epstein et al., 2018) and are useful for comparing reconstruction progress across the affected areas.

Increasing levels of aid receipt and additional household resource contributions increase the probability of a household taking reconstruction-related action. However, the marginal effect of cash support alone is relatively small (8%), suggesting that the first tranche of reconstruction funding or comparative I/NGO aid is not by itself a sufficient source of assistance for most households to begin reconstruction. These results hold across all survey months and all levels of additional resource contribution. This does not necessarily imply that the first tranche is not adequate financial support, but rather that financial support alone does not significantly increase the probability that a household starts rebuilding. It is also plausible that some households might refrain from starting reconstruction until they feel they have the resources (financial, technical, etc.) necessary to proceed towards the second and third tranche requirements which require adherence to specific seismically resistant building designs. The marked increase in probability of reconstruction action with receipt of an engineering consultation points toward this reality.

Another important finding is the asymmetric reconstruction probabilities across different levels of additional resource contribution. Across all aid levels, there are clear gaps between households with existing resources and those without. More importantly however, the relative differences in reconstruction probabilities increase as households receive more government aid. Based on 3.3, households with no or partial no responses for additional resource contribution have median probability increases of approximately 10-25% when comparing no aid to full aid, while households with partial yes or yes responses have increases of 50-70%. While these results are not surprising, they raise questions over how reconstruction aid should be allocated following major earthquakes. In Nepal, grant funding and other technical and social support mechanisms are intended to be provided equally to eligible households independent of socio-economic conditions (of Nepal, 2016). This study clearly shows that equal aid distribution does not correspond to equal reconstruction probabilities for households with different socio-economic conditions, at least for the first year of the reconstruction process. Understanding that equal aid distribution does not necessarily correspond to equitable recovery is important for communicating reconstruction progress. Although statistics on tranche disbursement are important for managing the reconstruction process, these findings suggest that they do not fully represent local recovery progress.

The Bayesian IRT modeling approach adopted here is a promising method for studying earthquake reconstruction progress. One of the key strengths of the IRT approach lies in estimating household-level latent reconstruction ability as a means to connect disparate survey responses. Reconstruction perception surveys are already collected regularly for monitoring reconstruction feedback, but have not previously been used in such an application. In this study, it is assumed that geographic location, household willingness to contribute resources to reconstruction, and aid receipt adequately predict latent reconstruction ability. These variables are supported by evidence from other longitudinal reconstruction Nepal (The Asia Foundation, 2017) but may not fully capture all of the factors relevant for household reconstruction. Regardless, compared to other types of models, the IRT approach offers the distinct advantage of modeling nonlinear responses between household ability and the probability of reconstruction with differentially functioning item-response curves. This is important for the reconstruction context because it is unlikely that rebuilding is equally difficult in all areas, even for households that have received similar levels of assistance. Using a Bayesian multi-level model to account for these differences models both the population-level estimate of reconstruction probabilities and regional deviations while directly quantifying the associated uncertainties. For the CFP survey data, this approach leverages the strength of three survey rounds to improve parameter estimates rather than treating each month as a separate dataset.

However, there are several limitations in the CFP survey data that informed model design and are worth mentioning for informing future survey design. The CFP surveys incorporated both reconstruction progress and housing damage state into a single question. Therefore, the response variable used in this study defines a positive response as any reported reconstruction progress (i.e. repairs or rebuilding), irrespective of the initial damage level. This is still a useful comparison, particularly in the context of evaluating early-phase reconstruction progress, but more detailed analysis could be performed if initial damage state data were collected for every household regardless of whether rebuilding had begun. Specifically, it would be useful to know whether a structure needed repairs or full reconstruction, given that the associated household's decision making almost certainly varies under these different circumstances. In the IRT framework implemented in this study, the effects of these differences get embedded into district-level estimates of discrimination and difficulty. This has the potential to bias results for isolated areas that were disproportionately damaged or undamaged relative to their surroundings.

The CFP surveys are also missing information on the timing of both aid receipt and reconstruction actions. Without timing on reconstruction actions, the model probabilities are relative to the timing of the surveys (11, 12, and 14 months after the post-disaster recovery framework was finalized). While the one-year time frame seems reasonable for assessing initial recovery progress, the reconstruction probabilities presented here are likely to change over time. Additionally, the current model has no way to distinguish between a household that received aid as soon as possible versus a household who received aid the day before being surveyed. In theory damage surveys and grant enrollment were complete before April 2017, but many households in the surveys report disbursement delays and ongoing redress of grievances. Some of these uncertainties might be ameliorated by the random effects components if households in the same areas received aid at similar times, but it is not clear the extent to which this is the case. In either case, all model parameters have full uncertainty estimates to account for these shortcomings.

To conclude, this study provides estimates of household-level reconstruction actions using Bayesian IRT modeling and data from Nepal's Inter-Agency CFP surveys. In translating household-survey data to population and district level estimates of reconstruction probabilities across varying ability levels, the model framework presented here represents a tangible step forward in bridging the gap between local qualitative case studies and coarse reconstruction statistics. The findings for Nepal suggest that households with existing resources are able to utilize the government-provided reconstruction assistance to a much greater degree than other households. It is hoped that elaborating these differences quantitatively serves as a guide for future research into recovery equity and provides an additional mechanism for evaluating earthquake reconstruction and recovery progress.

3.7 Acknowledgments

This research was performed under a National Science Foundation Graduate Research Fellowship, grant no. DGE-1450079.

3.8 Data Availability

The Inter-Agency Common Feedback Survey Data is available at: http://www.cfp.org.np

3.9 Works Cited

Alegana, V. A., Wright, J., Pezzulo, C., Tatem, A. J., and Atkinson, P. M. Treatment-seeking behaviour in low- and middle-income countries estimated using a Bayesian model. *BMC Medical Research Methodology*, 17(1), Dec. 2017. ISSN 1471-2288. doi: 10.1186/s12874-017-0346-0. URL http://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-017-0346-0.

Alegana, V. A., Maina, J., Ouma, P. O., Macharia, P. M., Wright, J., Atkinson, P. M., Okiro, E. A., Snow, R. W., and Tatem, A. J. National and sub-national variation in patterns of febrile case management in sub-Saharan Africa. *Nature Communications*, 9(1), Dec. 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-07536-9. URL http://www.nature.com/articles/s41467-018-07536-9.

Bank, W. Guide to Developing Disaster Recovery Frameworks: Sendai Conference Version. Technical report, World Bank, Washington, DC., Mar. 2015.

Bock, D. R. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, 37(1):29–51, Mar. 1972. ISSN 1860-0980. doi: 10.1007/BF02291411.

Bownas, R. and Bishokarma, R. Access after the earthquake: the micro politics of recovery and reconstruction in Sindhupalchok District, Nepal, with particular reference to caste. *Contemporary South Asia*, 0(0):1–17, Dec. 2018. ISSN 0958-4935. doi: 10.1080/09584935.2018.1559278. URL https://doi.org/10.1080/09584935.2018.1559278.

Brkner, P.-C. brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1):1–28, Aug. 2017. ISSN 1548-7660. doi: 10.18637/jss.v080.i01. URL https://www.jstatsoft.org/index.php/jss/article/view/v080i01.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. Stan: A Probabilistic Programming Language. *Journal of Statistical Software*, 76(1):1–32, Jan. 2017. ISSN 1548-7660. doi: 10.18637/jss.v076.i01. URL https://www.jstatsoft.org/index.php/jss/article/view/v076i01.

Chang, Y., Wilkinson, S., Potangaroa, R., and Seville, E. Interpreting Resourcing Bottlenecks of Post-Wenchuan Earthquake Reconstruction in China. *International Journal of Strategic Property Management*, 14(4):314–331, Dec. 2010. ISSN 1648-715X, 1648-9179. doi: 10.3846/ijspm.2010.24. URL https://journals.vgtu.lt/index.php/IJSPM/article/view/5739.

Chatterjee, R. and Okazaki, K. Household Livelihood Recovery after 2015 Nepal Earthquake in Informal Economy: Case Study of Shop Owners in Bungamati. *Procedia Engineering*, 212: 543–550, Jan. 2018. ISSN 1877-7058. doi: 10.1016/j.proeng.2018.01.070. URL http://www.sciencedirect.com/science/article/pii/S1877705818300882.

Comfort, L. K. and Joshi, J. Scalability and Sustainability in Uncertain Environments: Transition to Recovery from the 2015 Gorkha, Nepal, Earthquakes. *Earthquake Spectra*, 33(S1): S385–S401, Dec. 2017. ISSN 8755-2930. doi: 10.1193/113016EQS217M. URL http://earthquakespectra.org/doi/10.1193/113016EQS217M.

Daly, P., Ninglekhu, S., Hollenbach, P., Duyne Barenstein, J., and Nguyen, D. Situating local stakeholders within national disaster governance structures: rebuilding urban neighbourhoods following the 2015 Nepal earthquake. *Environment and Urbanization*, 29(2):403–424, Oct. 2017. ISSN 0956-2478, 1746-0301. doi: 10.1177/0956247817721403. URL http://journals.sagepub.com/doi/10.1177/0956247817721403.

Dunford, M. and Li, L. Earthquake reconstruction in Wenchuan: Assessing the state overall plan and addressing the forgotten phase. *Applied Geography*, 31(3):998–1009, July 2011. ISSN 0143-6228. doi: 10.1016/j.apgeog.2011.01.001. URL http://www.sciencedirect.com/science/article/pii/S0143622811000026.

Epstein, K., DiCarlo, J., Marsh, R., Adhikari, B., Paudel, D., Ray, I., and Mren, I. Recovery and adaptation after the 2015 Nepal earthquakes: a smallholder household perspective. *Ecology and Society*, 23(1), Feb. 2018. ISSN 1708-3087. doi: 10.5751/ES-09909-230129. URL https://www.ecologyandsociety.org/vol23/iss1/art29/.

Fox, J.-P. *Bayesian Item Response Modeling: Theory and Applications*. Statistics for Social and Behavioral Sciences. Springer-Verlag, New York, 2010. ISBN 978-1-4419-0741-7. URL https://www.springer.com/gp/book/9781441907417.

Fox, J.-P. and Glas, C. A. W. Bayesian estimation of a multilevel IRT model using gibbs sampling. *Psychometrika*, 66(2):271–288, June 2001. ISSN 0033-3123, 1860-0980. doi: 10.1007/BF02294839. URL http://link.springer.com/10.1007/BF02294839.

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., and Gelman, A. Visualization in Bayesian workflow. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 182(2): 389–402, Feb. 2019. ISSN 09641998. doi: 10.1111/rssa.12378. URL http://doi.wiley.com/10.1111/rssa.12378.

Gelman, A., Simpson, D., and Betancourt, M. The Prior Can Often Only Be Understood in the Context of the Likelihood. *Entropy*, 19(10):555, Oct. 2017. ISSN 1099-4300. doi: 10.3390/e19100555.

Glas, C. A. W. and van der Linden, W. J. Computerized Adaptive Testing With Item Cloning.

Applied Psychological Measurement, 27(4):247–261, July 2003. ISSN 0146-6216, 1552-3497. doi: 10.1177/0146621603027004001. URL http://journals.sagepub.com/doi/10.1177/0146621603027004001.

Hall, M., Lee, A., Cartwright, C., Marahatta, S., Karki, J., and Simkhada, P. The 2015 Nepal earthquake disaster: lessons learned one year on. *Public Health*, 145:39–44, Apr. 2017. ISSN 00333506. doi: 10.1016/j.puhe.2016.12.031. URL https://linkinghub.elsevier.com/retrieve/pii/S0033350616304528.

He, L., Aitchison, J. C., Hussey, K., Wei, Y., and Lo, A. Accumulation of vulnerabilities in the aftermath of the 2015 Nepal earthquake: Household displacement, livelihood changes and recovery challenges. *International Journal of Disaster Risk Reduction*, 31:68–75, Oct. 2018. ISSN 22124209. doi: 10.1016/j.ijdrr.2018.04.017. URL https://linkinghub.elsevier.com/retrieve/pii/S2212420918301079.

Mishra, A., Ghate, R., Maharjan, A., Gurung, J., Pathak, G., and Upraity, A. N. Building ex ante resilience of disaster-exposed mountain communities: Drawing insights from the Nepal earthquake recovery. *International Journal of Disaster Risk Reduction*, 22:167–178, June 2017. ISSN 2212-4209. doi: 10.1016/j.ijdrr.2017.03.008. URL http://www.sciencedirect.com/science/article/pii/S2212420916307245.

of Nepal, G. Nepal Earthquake 2015: Post Disaster Recovery Framework 2016-2020. Technical report, Government of Nepal, Kathmandu, Nepal, Apr. 2016.

Sulis, I. and Toland, M. D. Introduction to Multilevel Item Response Theory Analysis: Descriptive and Explanatory Models. *The Journal of Early Adolescence*, 37(1):85–128, Jan. 2017. ISSN 0272-4316, 1552-5449. doi: 10.1177/0272431616642328. URL http://journals.sagepub.com/doi/10.1177/0272431616642328.

Swaminathan, H. and Gifford, J. A. Bayesian estimation in the two-parameter logistic model. *Psychometrika*, 50(3):349–364, Sept. 1985. ISSN 1860-0980. doi: 10.1007/BF02294110. URL https://doi.org/10.1007/BF02294110.

Swaminathan, H. and Gifford, J. A. Bayesian estimation in the three-parameter logistic model. *Psychometrika*, 51(4):589–601, Dec. 1986. ISSN 1860-0980. doi: 10.1007/BF02295598. URL https://doi.org/10.1007/BF02295598.

The Asia Foundation. Independent Impacts and Recovery Monitoring Phase 4: April 2017 Synthesis Report. Technical report, The Asia Foundation, 2017.

Wu, J. Y. and Lindell, M. K. Housing Reconstruction After Two Major Earthquakes: The 1994 Northridge Earthquake in the United States and the 1999 Chi-Chi Earthquake in Taiwan. *Disasters*, 28(1):63–81, Mar. 2004. ISSN 03613666. doi: 10.1111/j.0361-3666.2004.00243.x. URL http://doi.wiley.com/10.1111/j.0361-3666.2004.00243.x.

3.10 Appendix

3.10.1 CFP Survey Data

Table 3.1 shows the survey questions used in the analysis and the coding scheme applied to the responses. Survey questions and possible responses are consistent across all three survey months. The aid and home status questions are coded into binary responses for suitability in the modeling framework. Willingness to commit additional resources is already structured as an ordinal variable and is not modified. Responses that include NAs (n=387) are left out of the analysis because they are ignored in Stan's No U-Turn Sampler algorithm. Each survey month originally included 2100 responses, leaving n=5913 complete responses for analysis.

Table 3.1: Inter-Agency Common Feedback Survey questions and corresponding codings used for analysis.

Question:	**Response Options:**	**Coding:**
What is the current status of your home?	Reconstruction Completed	Rebuilding
	Reconstruction Started	Rebuilding
	Minor Damaged	Not Rebuilt
	Heavily Damaged	Not Rebuilt
	Completely Destroyed (Rubble Cleared)	Not Rebuilt
	Completely Destroyed (Rubble Not Cleared)	Not Rebuilt
	Not Damaged	NA
What types of housing reconstruction support have you received?*	Direct Cash Support	Cash Yes/No
Have you consulted an engineer for your housing reconstruction needs?**	Have Consulted	Engineer Yes
	Have Not Consulted	Engineer No
	Plan to Consult	Engineer No
	Don't Plan to Consult	Engineer No
	Don't Know/Refused	NA
Have you been able to commit your own resources?	Completely Yes	Yes
	Somewhat Yes	Partial Yes
	Not Very Much	Partial No
	Not At All	No

*Only direct cash transfers (housing grant, I/NGO, etc.) were considered.
**The source of an engineering consultation was not considered.

3.10.2 Model Diagnostics

Prior Predictive Simulation

As mentioned in section 3.2.2, a prior predictive simulation draws MCMC samples directly from the joint prior distribution and therefore can be used to visually evaluate prior choices. This process is useful for understanding how the model behaves before any data is added. The objective here is to verify that draws from the prior predictive distribution represent data that could actually be observed given relevant domain knowledge, ideally with some mass for improbable data and no mass for impossible data (Gelman et al., 2017, Gabry et al., 2019). Figure

3.4 shows 100 simulated item-response curves drawn from the prior predictive distribution. The curve shapes reflect the prior choicespositive discrimination values and the lower-threshold favoring lower valueswhile allowing for a wide range of other possibilities. Although nearly flat or nearly step-function curve shapes are implausible, they are permitted here because they are technically possible given domain knowledge.

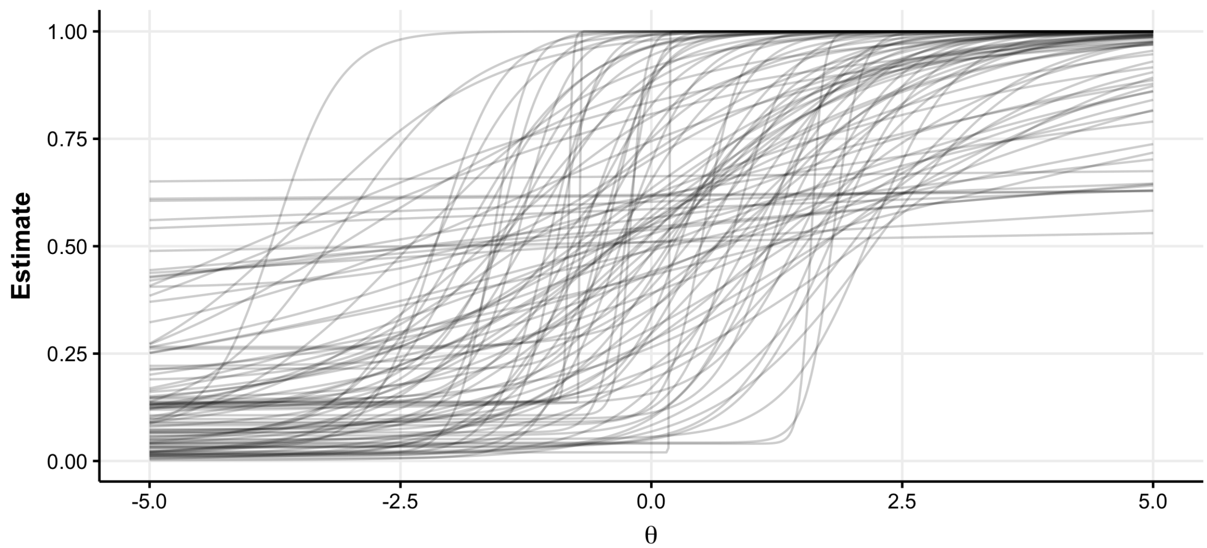

Figure 3.4: 100 simulated draws from the prior predictive distribution.

Model Parameters and Convergence Diagnostics

Table 3.2 provides posterior parameter estimates, confidence intervals, and convergence statistics for the model. Considering 4,000 post-warmup iterations, effective sample sizes for all parameters are well above levels of concern. $\hat{R}$ values indicate convergence for all parameters. MCMC trace plots and densities for each parameter are shown in Figures 3.5 and 3.6. The trace plots appear well mixed and do not indicate any obvious pathologies. Model computation in Stan did not produce any divergent transitions or other runtime warnings.

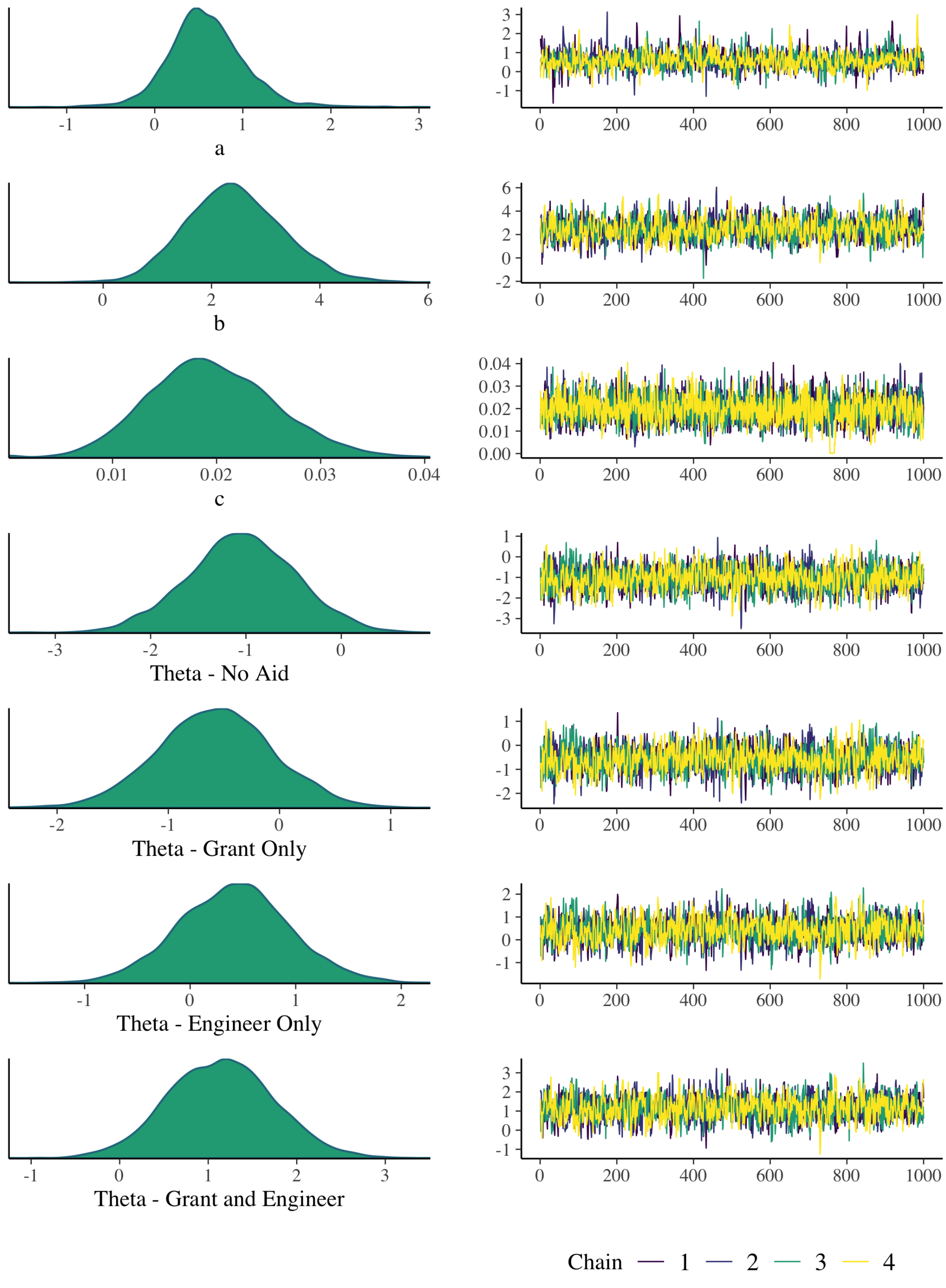

Figure 3.5: MCMC parameter density estimates and trace plots.

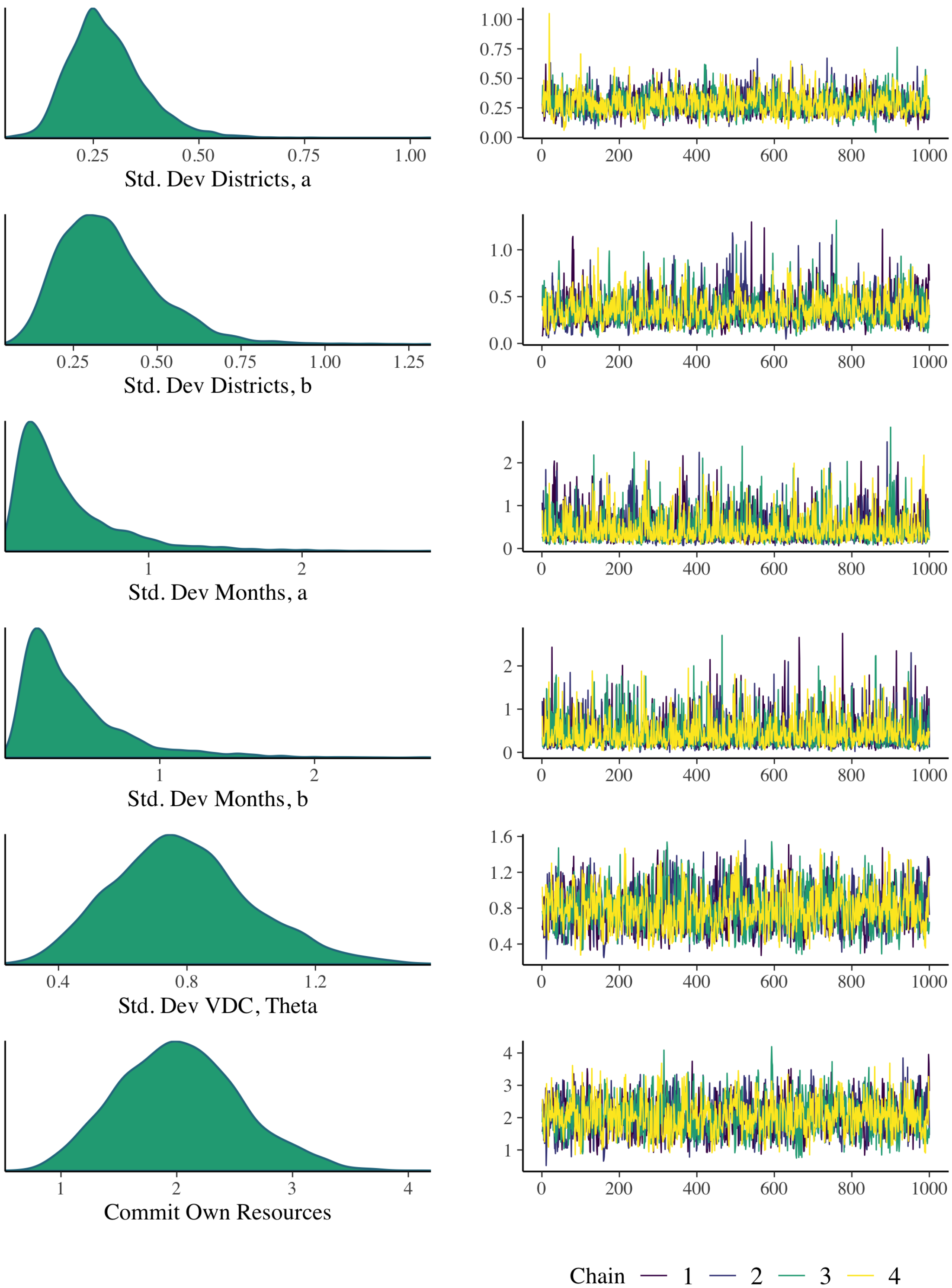

Figure 3.6: MCMC parameter density estimates and trace plots.

Table 3.2: Posterior parameter estimates.

Group Level Effects:					
Month (n=3)	Estimate	Est. Error	95% CI	Eff. Sample	$\hat{R}$
sd(a)	0.45	0.33	(0.12 - 1.40)	2086	1.00
sd(b)	0.43	0.34	(0.08 - 1.38)	2554	1.00
District (n=14)					
sd(a)	0.28	0.09	(0.14 - 0.48)	1398	1.00
sd(b)	0.36	0.15	(0.14 - 0.71)	1232	1.00
VDC (n=245)					
sd(θ)	0.79	0.21	(0.41- 1.25)	1318	1.00
Population Level Effects:					
a	0.59	0.44	(-0.23 - 1.48)	1340	1.00
b	2.43	0.89	(0.83 - 4.27)	1320	1.00
c	0.02	0.01	(0.01 - 0.03)	2016	1.00
θ_{NoAid}	-1.04	0.57	(-2.18 - 0.06)	2120	1.00
$\theta_{GrantOnly}$	-0.58	0.52	(-1.59 - 0.42)	2099	1.00
$\theta_{Engineer}$	0.42	0.52	(-0.60 - 1.47)	1795	1.00
$\theta_{GrantEngineer}$	1.15	0.60	(-0.01 - 2.33)	1507	1.00
$\theta_{CommitOwnResources}$	2.03	0.53	(1.08 - 3.13)	1494	1.00
Simplex Parameters*:					
$\theta_{Commit-No,PartialNo}$	0.22	0.06	(0.11 - 0.34)	4196	1.00
$\theta_{Commit-PartialNo,PartialYes}$	0.55	0.06	(0.44 - 0.68)	3327	1.00
$\theta_{Commit-PartialYes,Yes}$	0.23	0.06	(0.10 - 0.34)	3188	1.00

*Simplex parameters specify the differences between adjacent categories

3.10.3 Model Validation

Posterior predictive checks compare data generated from the posterior distribution to a set of observed data to evaluate model fit. A well calibrated model will generate data that resembles observed data and performs consistently across grouping levels (Gabry et al. 2019). To evaluate model performance, a random 25 percent sample (n=1478) of the data are reserved as an independent test set and the model is fit to the remaining 75 percent (n=4435). Reconstruction outcomes are then estimated from the posterior distribution and compared to the observed data.

Figure 3.7a show 100 predicted draws in black compared to the observed data in red for the entire test set. The same predictions are separated out by predictor variable in Figures 3.7b and 3.7c, and districts and months in Figures 3.8 and 3.9.

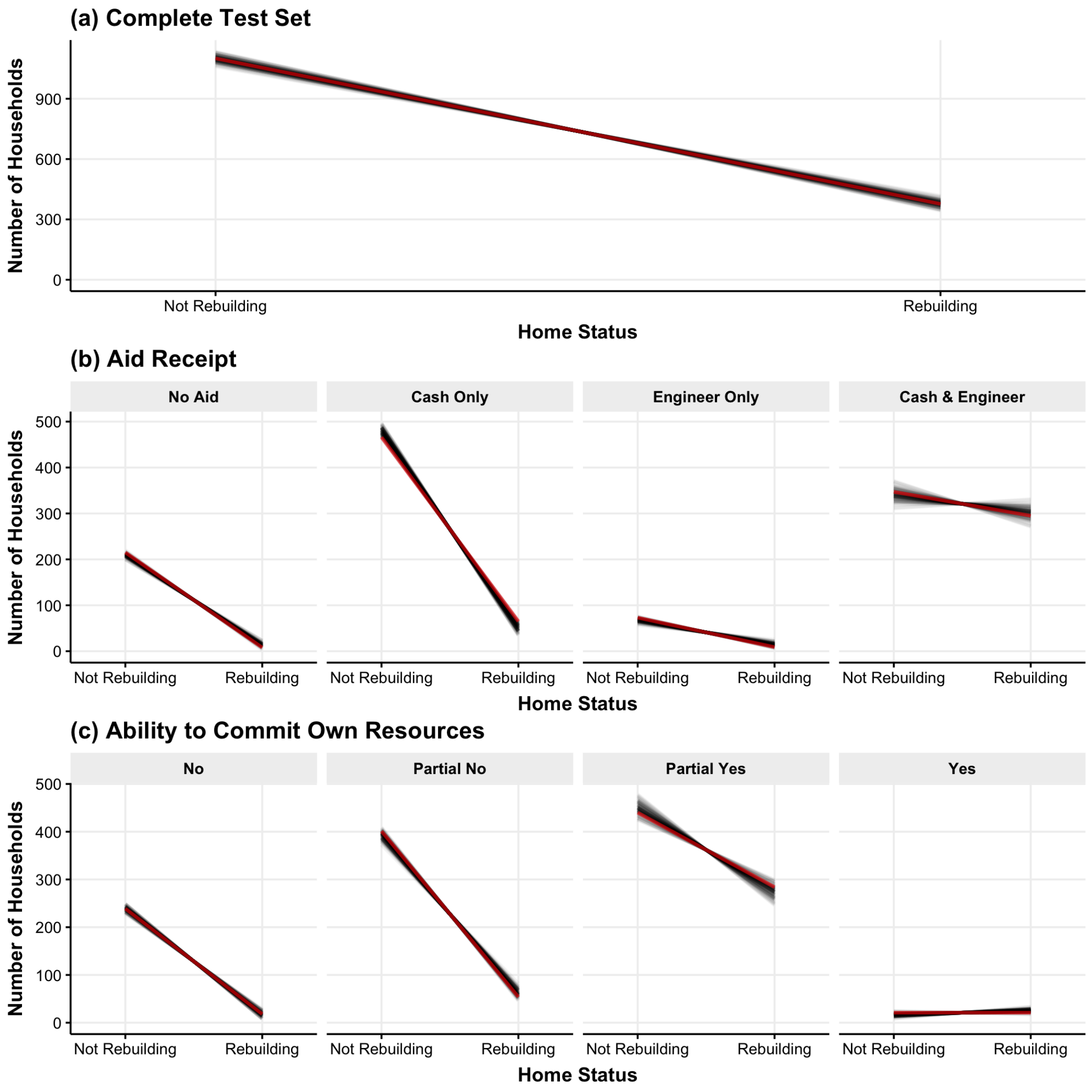

Figure 3.7: 100 draws from the posterior predictive distribution compared to observed data. (a) Results for the entire test set, (b) Results separated by aid receipt variable, (c) Results separated by ability to commit own resources variable.

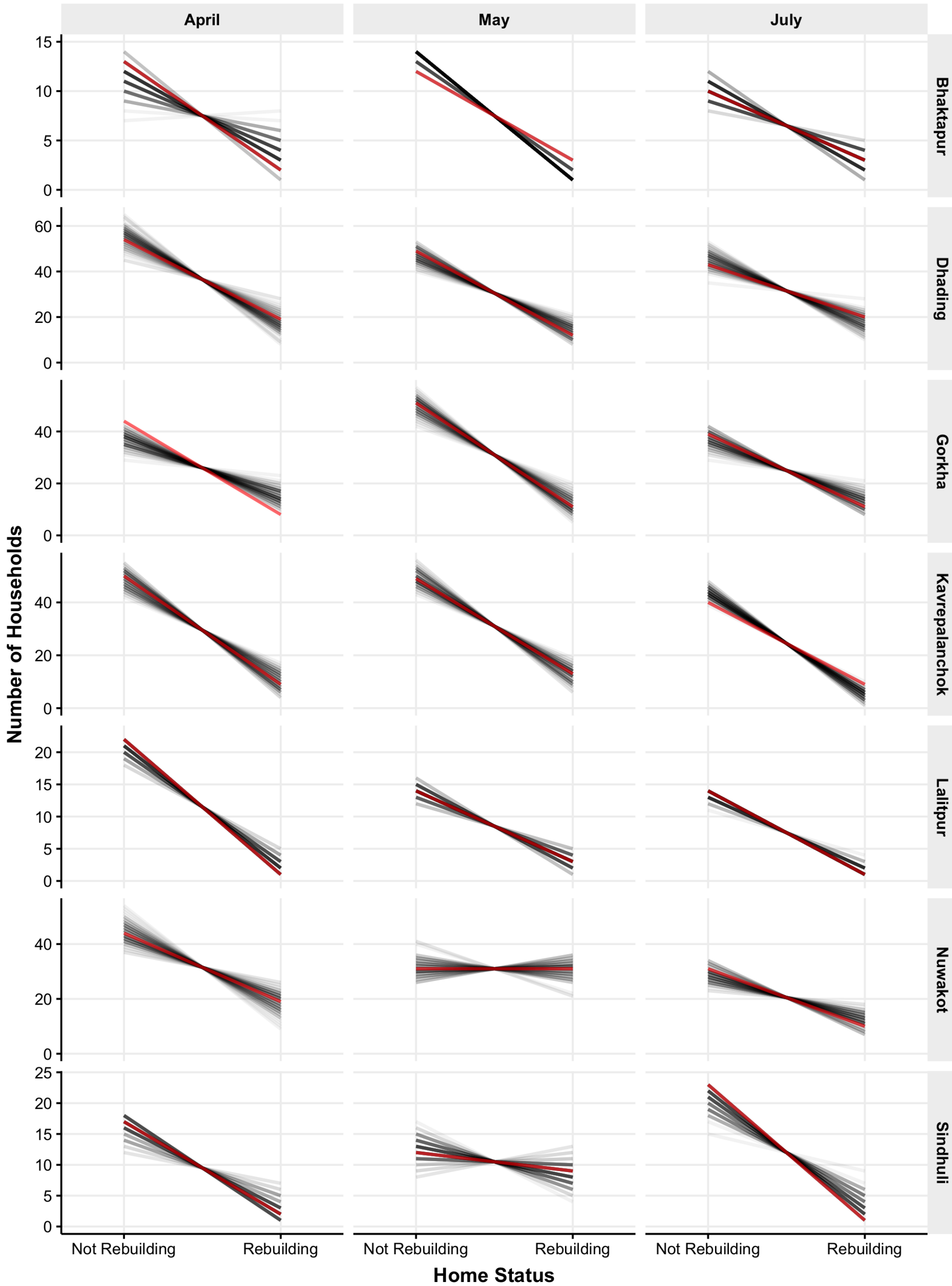

Figure 3.8: 100 draws from the posterior predictive distribution compared to observed data, separated by survey month and district.

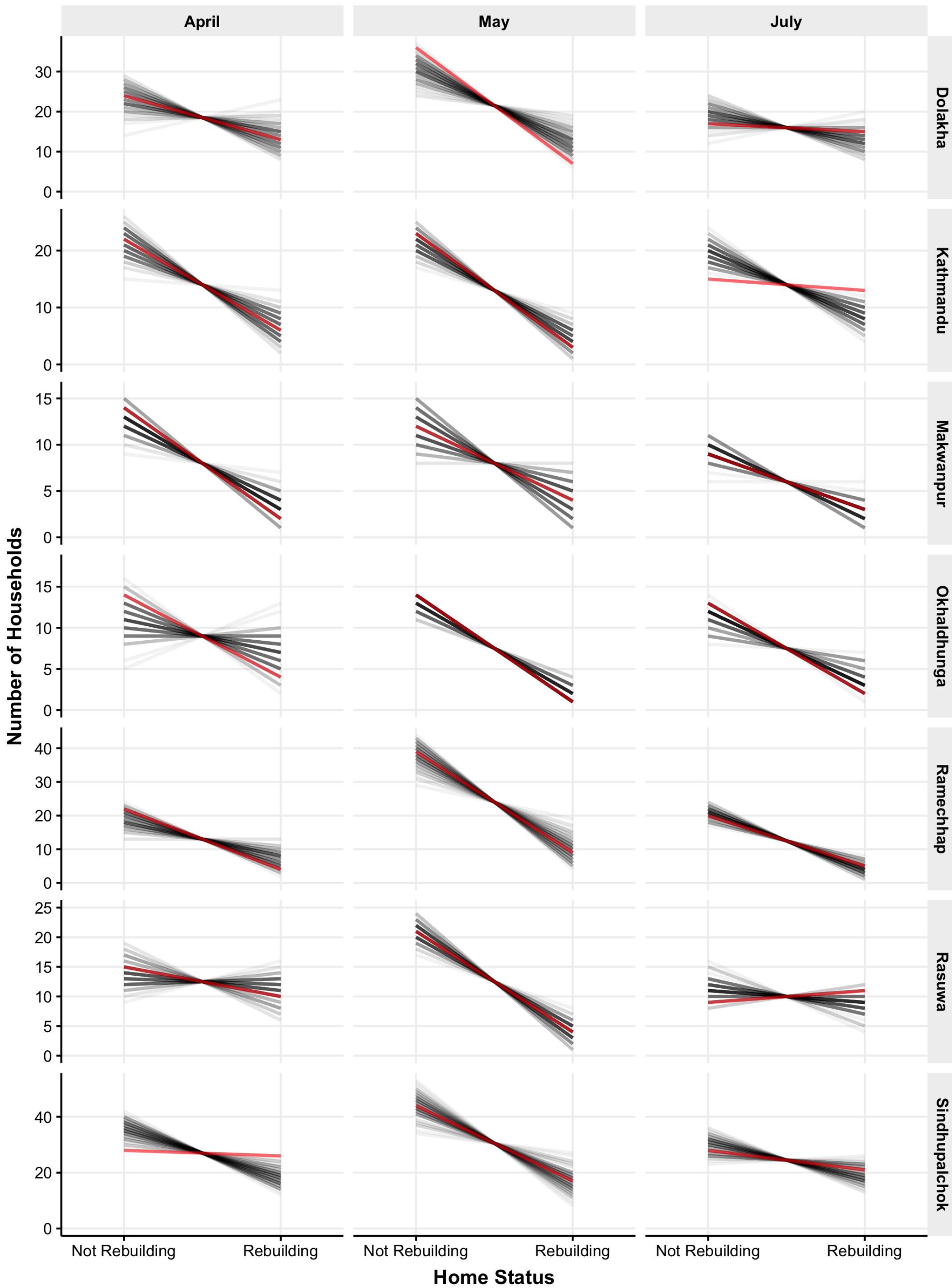

Figure 3.9: 100 draws from the posterior predictive distribution compared to observed data, separated by survey month and district.

4: A SPATIAL INTERPOLATION MODEL FOR HIGH-RESOLUTION MAPPING OF EARTHQUAKE DAMAGES FROM GEOLOCATED CLUSTER DATA

This chapter corresponds to the following in-preparation paper: Wilson, B.S. (2019). A spatial interpolation model for high-resolution mapping of earthquake damages from geolocated cluster data.

4.1 Abstract

Modeled damage estimates are an important source of information in the hours to weeks following major earthquake disasters, but often lack sufficient spatial resolution for highlighting specific areas of need. Using damage assessment data from the 2015 Gorkha, Nepal Earthquake, this paper evaluates a spatial regression model for interpolating geolocated damage survey data onto a 1 x 1 km^2 grid. The proposed approach uses a combination of geospatial covariate data and Gaussian spatial process random effects modeling to estimate the percentage of structures attaining complete damage states from sparse survey clusters using the INLA-SPDE method. Model performance is evaluated across fifty iterations of 100, 250, and 1000 simulated survey clusters and compared to observed damage assessments and model predictions using more traditional fragility-based methods. Results show strong model fit to observed values, with mean absolute errors of .17, .13, and .11 and correlation coefficients of .75, .82, and .85 for increasing numbers of survey clusters. These results show significant improvements over existing methods with a fraction of the damage surveys that were available within several weeks after the Gorkha event. Thus, with sufficient rapid damage assessment mobilization, the proposed spatial modeling framework offers improved damage estimates and higher spatial resolution while remaining

tractable within the time frame required to deliver a Post Disaster Needs Assessment.

4.2 Introduction

In the hours to weeks following major earthquake disasters, detailed information on the spatial variability, extent, and severity of damages is often sparse. This lack of consistent and verifiable post-disaster information poses major challenges for rapid emergency response efforts (Comfort et al., 2005, Goodchild and Glennon, 2010, Lallemant et al., 2017). Until field surveyors can be mobilized at scale, disaster response decisions are informed by coarse modeled damage estimates, scattered eye-witness reports, and any remotely-sensed damage assessments that might be available (Goodchild and Glennon, 2010, Xie et al., 2016, Lallemant et al., 2017). However, it is not well understood how these disparate sources of information should be synthesized, in part due to significant scale differences between event-scale impact assessments produced by systems like the U.S. Geological Survey's Prompt Assessment of Global Earthquakes for Response (PAGER) and more localized sources of information. Understanding how to effectively model post-disaster damages has important applications for informing post-disaster needs assessments and requests for international development aid.

Common model-based approaches for estimating earthquake damages rely on building stock records and functions that describe the probability of damage at various shaking levels for different housing typologies (Whitman et al., 1997, Yeh et al., 2006, Kircher et al., 2006, Robinson et al., 2018). These relationships can be derived empirically, semi-empirically, or analytically depending on the availability of relevant local seismological studies (Jaiswal et al., 2011, Porter, 2014). While modeled areal damage estimates are successfully applied for rapid order-of-magnitude impact estimates in systems like PAGER, their accuracy at high spatial resolutions is strongly dependent on the quality of input data (Erdik et al., 2011, Jaiswal et al., 2011, Lallemant et al., 2017). On the other hand, field-based engineering assessments provide detailed ground-truthed damage data, but are relatively sparse in the weeks following a major earthquake. Over 60,000 rapid visual assessments were collected in several weeks following the 2015 Gorkha, Nepal Earthquake to inform the post-disaster needs assessment (PDNA), yet still represented less than 10% of affected structures. As a result, the damage statistics that inform requests for disaster aid are often based on modeled estimates. The timeline for delivery of a PDNA to funding stakeholders (approximately a month) is simply too short to fully survey the damages (Lallemant et al., 2017).

With significant focus in the Sendai Framework for Disaster Risk Reduction placed on leaving no one behind and accounting for vulnerable populations, improving the capacity of damage models to capture spatial heterogeneities is an important goal. Accordingly, this paper develops a spatial interpolation damage modeling framework that uses geolocated clusters of surveyed damages and readily available geospatial covariates to estimate the probability of complete structural damage on a high resolution grid. The proposed framework is explicitly designed designed to be implementable within the timeline for PDNA delivery, only requiring a U.S. Geological Survey Shakemap, readily available gridded covariate layers, and sparse clusters of rapid damage assessments that could be feasibly collected in several weeks. The model is developed using the Integrated Nested Laplace Approximation (INLA) and Stochastic Partial Differential Equation (SPDE) approaches. The INLA methodology efficiently implement approximate Bayesian inference for a subset of models that can be defined with latent Guassian Markov random fields (Rue et al., 2009). The SPDE approach provides a computationally convenient way to implement models with spatial effects using INLA methods (Lindgren et al., 2011). Similar spatial interpolation models have been applied to a wide variety of development and demographic contexts (Tatem et al., 2014, Bhatt et al., 2015, Bosco et al., 2017, Utazi et al., 2018b), leveraging the correlations of particular response variables with geographic, environmental, or socio-demographic variables for which higher resolution data is available.

This study uses data from the Gorkha Earthquake to illustrate the proposed approach, an event for which field surveyed post-earthquake damage state designations are available. The rest of the paper is organized as follows. Section 2 reviews the fragility-based damage modeling approaches used in many applications. This section provides relevant background for understanding the damages statistics in the Gorkha Earthquake PDNA and motivates the proposed improvements. Section 3 covers the input data sources used in the model and their relevance to damage estimation. The details of the INLA-SPDE approach and model evaluation are described in section 4, with model results for different numbers of simulated survey clusters presented in section 5. The final section discusses model performance, limitations, and some recommendations for future work.

4.3 Motivation

Earthquake-related building damages are a function of both shaking intensity at a given location and the seismic resistance of exposed structures. Fragility curves capture this relationship, specifying the probability of of a structure exceeding a certain damage state at a given shaking level (Porter, 2014). A standard approach for estimating damages combines fragility curves, ground motions, and data on housing type distributions to estimate the percentages of structures attaining specific damage states for a given area (Whitman et al., 1997, Yeh et al., 2006, Kircher et al., 2006, Robinson et al., 2018). The damage estimates included in Nepal's PDNA used this type of methodology, drawing average Modified Mercalli Intensity values from the most recent USGS Shakemap, housing counts for four different building typologies from Nepal's 2011 Housing Census, and fragility curves from Guragain (2015) (of Nepal, 2015).

A slightly modified version of the PDNA estimate is implemented here to illustrate the performance for the 2015 Gorkha event. The methodological details provided in the PDNA are not sufficient to reproduce the exact estimates. Consequently, instead of estimating both fully damaged and partially damaged structures, only the most severe damage state is estimated (also termed 'complete'). This is assumed to be similar to the 'fully damaged' classification in the PDNA and is the quantity of interest for the spatial interpolation model. Nepal-specific fragility curves from Guragain (2015) are used for stone and brick buildings with mud or concrete mortar, the predominate building types in rural areas. For concrete and wood buildings, fragility curves are drawn from HAZUS-MH (FEMA, 2013). Ground motion and housing typology data are drawn from the same sources as the PDNA, a USGS Shakemap and the 2011 Housing Census, respectively. The PDNA estimates were calculated at the district level, but are reproduced here for village development committees (VDCs)—the finest spatial unit for which housing typology data is available. It is not clear why the PDNA used district-level estimates when VDC level data was available.

Figure 4.1 maps the modeled and observed damage levels across the eleven most-affected districts. The model predictions severely overestimate the percentage of completely damaged structures, showing nearly 90-100% damage rates across significant portions of the affected area. The mean absolute error (MAE) and root mean squared error (RMSE) across all VDCs are 42% and 52%, respectively, suggesting that traditional stone and mud structures performed better than estimated in many VDCs. This margin of error might be acceptable for an event-level 'or-

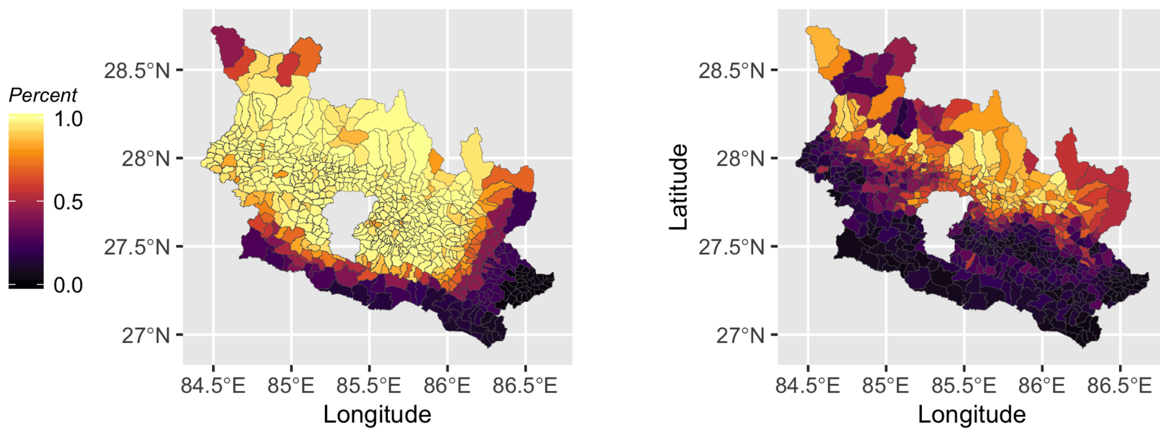

Figure 4.1: Modeled damage estimates using fragility curves and aggregated ground motions at the village development committee level (left) compared to observed damage levels (right).

der of magnitude' estimate produced within minutes to hours of an earthquake, but it is not clear whether these modeled estimates should be considered accurate enough to inform aid allocation decisions. Any modeling approach that does not have the ability to update model predictions with ground-truth damage assessments is tethered to the accuracy of uncertain ground motions and fragility curves.

4.4 Materials & Methods

4.4.1 The INLA-SPDE Approach

This study proposes an alternative damage modeling framework that integrates randomly sampled post-earthquake engineering damage assessments with gridded geospatial covariates to predict earthquake damages on a uniform grid. Consider a study region $A \in R^2$ discretized into a uniformly spaced grid with n_p grid points $s_1, ..., s_{n_p}$. For any given location i in the study region, let Y_i represent the number of structures at a specified damage state given N_i total structures. In this application, quantities Y_i and N_i are assumed to be observed for a given set of survey clusters

n_c with known locations and otherwise unknown at the grid point level. To estimate probabilities p_i at the grid point level, the model is defined by:

$$
\begin{aligned}
& Y_i|\mathbf{F_i}, \alpha, x_i, \theta \sim Binomial(N_i, p_i) \quad i = 1, ..., n_c + n_p \\
& logit(p_i) = \alpha + f(\mathbf{F_i}) + x_i \\
& \mathbf{x} \sim GF(0, \Sigma)
\end{aligned}
\tag{4.1}
$$

where $\mathbf{F_i}$ is the vector of covariate values, θ is a vector of hyperparameters, and $\mathbf{x}$ is the spatial latent Gaussian field that defines spatial random effects x_i. In this application, spatial random effects Σ are specified with a Matérn covariance function defined by a smoothness parameter v, scaling parameter k, and marginal variance σ_η^2 (Lindgren et al., 2011). A single latent Gaussian field is used for both the observed point-referenced data and the estimations on the prediction grid. Covariate values are defined across the survey area such that values exist for survey and prediction points.

Approximate Bayesian inference via the INLA-SPDE approach is used to fit the model in Equation 1.1) Rue et al. (2009). Principally, the INLA approach is an computationally efficient alternative to Markov chain Monte Carlo methods that produces numerical approximations of posterior parameters. For the vector of model hyperparameters $\theta = (\mathbf{F_i}, k, \sigma_\eta^2)$, the joint posterior is given by:

$$
\prod_{i=1}^{n_c+n_p} Binomial(Y_i; N_i, pi) \times N(\eta; 0, \Sigma) \times p(\theta) \tag{4.2}
$$

where $p(\theta)$ is the joint prior distribution of model parameters. The SPDE methodology

(Lindgren et al., 2011) is used for estimating the Gaussian field (η). This approaches defines a triangular mesh across the study region A using basis functions that provide a sparse representation of a Gaussian field with Matérn covariance at each triangulation node. A projector matrix is then used to linearly map values from triangulation nodes to points of interest inside the mesh. For full details on the SPDE approach, readers are referred to (Lindgren et al., 2011).

4.4.2 Geospatial Covariates

A set of gridded geospatial covariates are used to guide interpolation between damage levels at observed survey clusters. Covariates are chosen to capture ground motions and the spatial variability of housing types with different fragilities. Ground motions are included via a U.S. Geological Survey (USGS) Shakemap for Modified Mercalli Intensity. USGS Shakemaps are produced globally in near-real-time for all major earthquakes. Variability in housing types is reflected across three covariate layers: Shuttle Radar Topography Mission-derived elevation, Visible Infrared Imaging Radiometer Suite (VIIRS) nighttime lights, and distance to Open-Street Map major roads. These data are freely available from the WorldPop Project (WorldPop, 2018) and were selected based on the basis of domain knowledge about social, structural, and physiographic variation in Nepal (Muzzini and Aparicio, 2013, Chaulagain et al., 2015, Gautam and Chaulagain, 2016, Robinson et al., 2018). All Worldpop covariate values are re-sampled onto 1 x 1 km^2 grid cells to match the resolution of the USGS Shakemap.

VIIRS nighttime lights data is used as a proxy for urbanization. A majority of engineered structures and other reinforced concrete buildings in Nepal are located in urbanized areas (Gautam et al., 2016, Gautam and Chaulagain, 2016). Although population density could also be used to similar effect, nighttime lights provides a more direct signal for urbanization of the built environment. Distance to major roads is used as a proxy for remoteness and access to non-local building materials. Large portions of rural Nepal are only accessible on foot and are restricted in the types of building materials available (Muzzini and Aparicio, 2013). Most houses that are far from roads are traditionally constructed with stones and mud mortar. Elevation captures two potentially important variations in building materials. At coarse scales, elevation reflects differences between the predominate building types in the low-lying Terai region (wood and bamboo) versus the mid-hills region (stone with mud/cement mortar). At finer scales, there is a general tendency for reinforced concrete and other engineer structures to be located at lower relative elevations than more traditional construction built into terraced hillslopes. Together, these three covariates capture the dominant features distinguishing housing typologies in Nepal.

4.4.3 Post-Earthquake Damage Estimation from Simulated Survey Clusters

As previously mentioned, in the month following the Gorkha earthquake over 60,000 rapid visual building inspections were performed by volunteer engineers and architects trained by Nepal's National Society for Earthquake Technology (NSET) (Lallemant et al., 2017). While these original data are not openly available, a similar type of dataset can be simulated from damage surveys collected further along in the reconstruction process. In this paper, the proposed model is evaluated with simulated datasets containing increasing numbers of surveyed damage clusters (100, 500, and 1000) drawn from a complete set of damage assessments collected for all households across eleven rural districts as part of Nepal's Household Registration for Housing Reconstruction program. While detailed engineering assessments typically assign ordinal damage grades (Lallemant and Kiremidjian, 2015), PDNAs use a less detailed partial-complete damage spectrum (of Nepal, 2015). Accordingly, this study focuses only on predicting 'complete damage

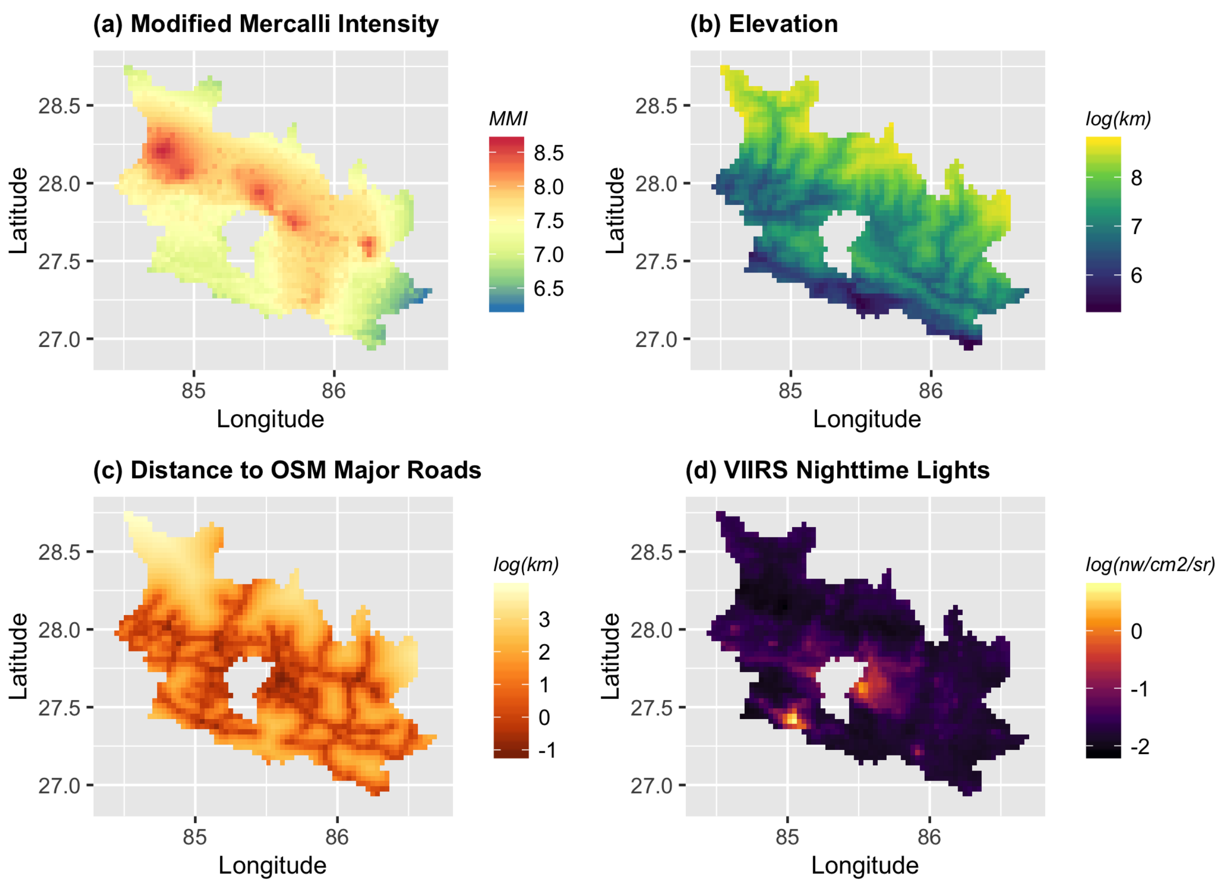

Figure 4.2: The four gridded covariates used in the spatial interpolation model: (a) Modified Mercalli Intensity, (b) elevation, (c) distance to OSM major roads, (d) VIIRS Nighttime Lights.

states', assumed to be those structures assigned the maximum possible damage grade (5 out of 5) in the existing surveys. These structures are priorities for emergency response and often require full reconstruction as opposed to repairs.

The cluster simulation procedure is described as follows. Cluster centroids are first drawn as random samples of all household locations in the study area. Each cluster is then assigned a fixed buffer of 250 meters and all households within the buffered zone are assumed to be surveyed irrespective of damage level. The decision to use a 250 meter buffer distance is somewhat arbitrary, but is chosen to be small enough such that a cluster could be surveyed by a small team of engineers reasonably quickly. Other sized clusters or even variable size clusters are usable as long as every structure in a cluster is surveyed. 1,000 surveyed clusters at 250 meters each includes approximately 55,000 households—close to the number of rapid damage assessments that were actually collected in a few weeks following the earthquake. 500 and 100 surveyed clusters cover approximately 28,000 and 5,500 households on average, respectively.

An iterative model fitting procedure with fifty instances for each number of survey clusters is implemented using the R-INLA package in R (Lindgren and Rue, 2015). Iteratively fitting the model to different randomly sampled cluster locations averages out the bias associated with predictions at any single cluster location to give a stronger overall picture of average model performance. For each iterations, a cluster set is generated and covariate values are extracted at the centroid locations. These data are used to fit Equation (1) with a SPDE triangular mesh constructed as a convex hull around the study area (see details in Appendix). Penalized complexity priors with the range set to the median distance between grid points are specified on the spatial random field (Simpson et al., 2017, Fuglstad et al., 2019) and weakly informative normal priors ($\sim N(0, 1e5)$ are placed on the covariate coefficients. Predictive performance is evaluated across all iterations for each number of survey clusters with a combination of three metrics: root mean square error, mean absolute error, correlations between observations and predictions at the grid level. Spatial plots for the median posterior estimates, standard deviations, and differences between predicted and observed values are also included.

4.5 Results

Figure 4.3A-E shows the median posterior estimates and standard deviations for the percentage of buildings sustaining complete damage in each 1 x 1 km^2 grid cell for 250, 500, and 1000 clusters. The corresponding parameter estimates are included in Table 1 in the Appendix. Damage patterns are positively correlated with MMI and distance to roads and negatively correlated with VIIRS nighttime lights, supporting prior assumptions for the covariates. Elevation switches from positive to negative correlations at 500 grid clusters, likely reflecting the lack of damage in the high mountain regions with no surveyed structures. Among included variables, shaking intensity has the strongest effect on modeled damage percentages. The estimates for the spatial range parameter vary between 1.18 and 1.48 (corresponding to an approximate distance of 112 to 152 kilometers), indicating strong spatial dependence in damage levels.

As expected, increasing the number of survey clusters increases the level of spatial heterogeneity in the modeled estimates and decreases the estimated uncertainties. Moving from 100 to 500 clusters has a larger impact on model estimates than the corresponding increase from 500 to 1000 clusters. This is predominately seen as reduction in uncertainty rather than a significant change among estimated damage levels. Uncertainties remain high in the Himalayas where few if any households are located. Compared to the areal damage model (Figure 4.1), all three cluster models more accurately capture the lower percentages of complete damage at moderate latitudes.

However, it is worth noting that the model predications and uncertainties from any single scenario will be affected to some degree by the survey cluster locations, with higher uncertainties further away from surveyed clusters. Optimizing cluster placement for interpolation accuracy is beyond the scope of this study, but could be a worthwhile direction in future work.

The mean absolute errors for the 100, 500, and 1000 cluster models are .17, .13, and .11, respectively (.23, .18, .15 RMSE). Correlations between predicated and observed damages (see Figure 4.6 in Appendix) are .71, .82 and .85. The MAE values imply an average difference between observed and predicted values at the grid cell level of 11-17%—a 25%+ improvement over the areal model with improved spatial resolution. As mentioned previously, model performance improves with additional survey clusters but reasonably accurate predictions can be obtained with a modest number of surveyed clusters. The average number of households included amongst the simulated 100 cluster models is less than 10% of the actual number of households surveyed following the Gorkha Earthquake.

Figure 4.4 shows the percent differences between predicted and observed values at the grid cell level. On the whole, all three models do a reasonable job of reproducing the observed damage levels without any major biases. However, the 100 cluster model shows more correlated errors compared to the 500 and 1000 cluster models. Including additional geospatial covariates or altering the included covariates could potentially improve model fit, although errors may also stem from ground motion uncertainties. USGS Shakemaps are themselves a modeled data product and often interpolate across large areas with no strong motion observations. Figure 4.4 also shows that all thee models contain cells where estimates are off by a significant margin (50-80%). These prediction errors are concentrated in cells that neighbor survey observations but have large differences in damage levels. The spatial random field favors smooth transitions between obser-

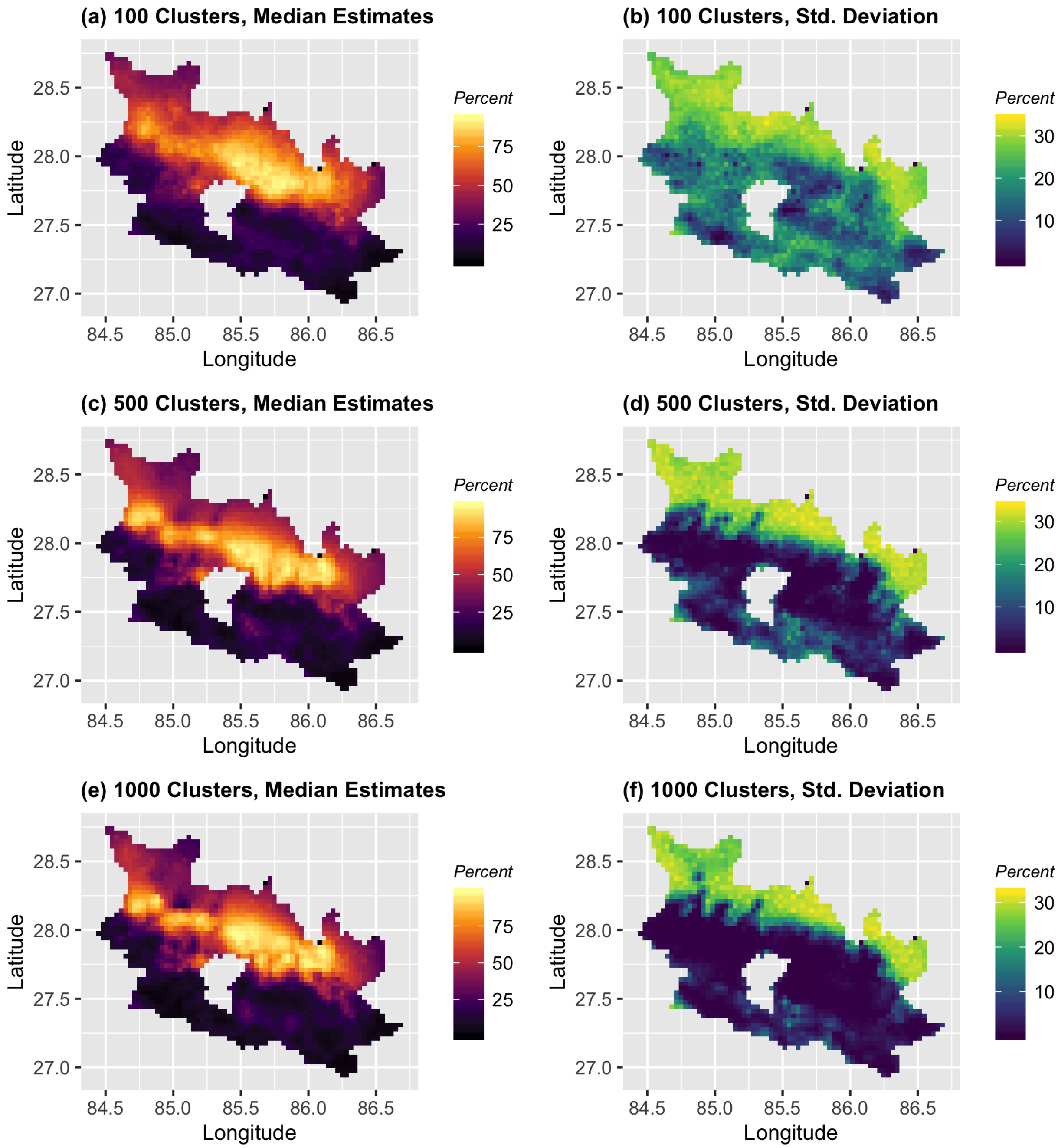

Figure 4.3: Median posterior estimates and standard deviations for the probability of complete damage at the grid cell level across fifty simulations using 100 (a-b), 500 (c-d), and 1000 (e-f) survey clusters.

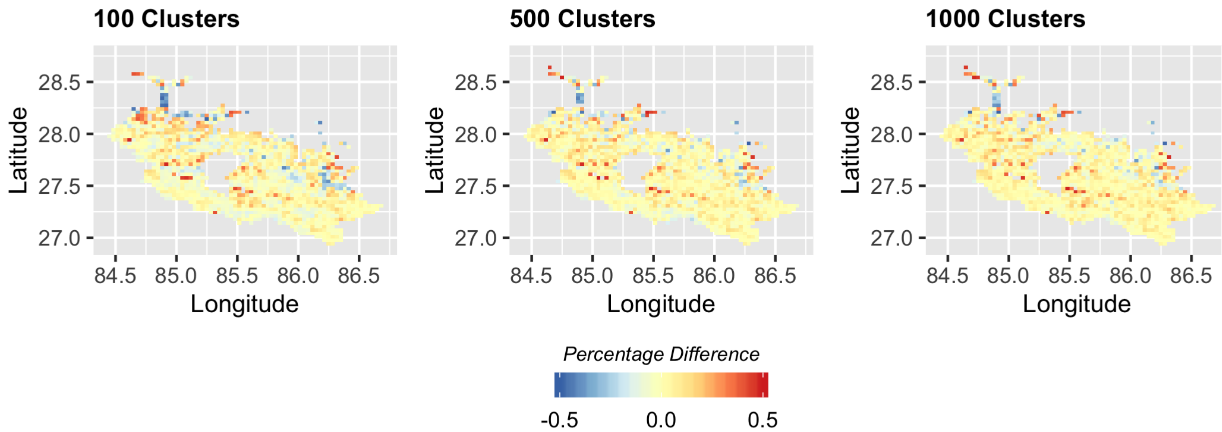

Figure 4.4: Median percentage difference between predicted and observed complete damage states at the grid cell level for 100, 500, and 1000 survey clusters (left to right).

vations and is not well suited for capturing rapid changes between neighboring pixels. Similar issues may also occur at sharp boundaries in the covariate layers. Although major prediction errors only exist in a small fraction of total model cells, they should be taken into account when interpreting results. Pixel to pixel differences are generally reliable, but the proposed model should not be used for targeting individual cells.

4.6 Discussion & Conclusion

Providing accurate models of earthquake damages on short-term timelines is a critical element for both disaster response and recovery planning. Requests for international aid following major earthquakes rely on the best damage assessments available approximately one month post-event. It is common for rapid damage assessments to be collected at varying locations across impacted areas to inform a post disaster needs assessment, but these data are not currently used in damage models due to their sparse distribution. This study developed a spatial interpolation damage modeling framework for estimating complete damage states on a high resolution prediction grid from this type of geolocated cluster-level data. The proposed model is based in linking surveyed damage states to gridded covariates that correlate with ground motions and building fragilities. The INLA-SPDE approach is used to fit the model in a computationally efficient manner. Results show improved performance over existing damage modeling approaches with as few as 100 clusters or approximately 5,500 surveyed structures. However, the true value of a spatial interpolation framework becomes apparent with closer to 500 or 1000 clusters (28,000 - 55,000 surveyed structures) where predictions capture detailed spatial heterogeneities and correlations between observed and predicted damage maps reach 85%.

The biggest advantage of a spatial interpolation approach compared to other damage models is the use of ground-truthed data. Damage models based using only ground motions and fragility curves are completely reliant on the accuracy of such data. In some regions this information is well constrained and analytical approaches can be an effective approach for estimating structural damages. However, these methods do not generalize particularly well to areas like Nepal with poor seismic instrumentation for capturing ground motions and a lack of detailed earthquake-engineering studies. Section 1.2 showed the large discrepancies between current models and observed damaged states for the Gorkha Earthquake, with average prediction errors over 40% (Figure 4.1). Using observed cluster-level damage states as a starting point circumvents this issue by making fewer assumptions on the a-priori relationship between ground motions and specific housing types. Rather, a more general statistical relationship is derived from observed damage levels across different local geographies with a spatial random field term capturing spatially correlated errors. In essence, the spatial interpolation model trades off an unknown assumption on fragilities for a known assumption on the distribution of different structural types across landscape characteristics. This particular approach works well in places like Nepal with strong socio-environmental variation among housing types.

Several limitations exist in the proposed spatial interpolation model that are worth mentioning. As described in section 1.3.2, the covariates used in the model are based on domain knowledge about the spatial variability of housing types in Nepal. In this application, Nepal benefits from a strong urban-rural divide that translates into roughly homogeneous housing types. Although other studies (e.g. Chaulagain et al., 2016, Robinson et al., 2018) use anywhere from 4-7 housing categories, the differences in associated fragilities between many of the included typologies is quite small. Hence, focusing on distinguishing between reinforced concrete, wood, and various stone/mud structures captures most of the fragility variability in Nepal at a one kilometer pixel resolution. Similar assumptions may not be reasonable in regions with more heterogeneous housing types, less spatial consistency among housing types, or weaker correlations between housing types and physiographic variables. The covariates used in this study are likely to be applicable in other seismically active mountainous countries, but the relationships should be thoroughly evaluated before generalizing this particular model to another context.

This modeling framework also utilized a comprehensive set of damage assessment data collected after a major earthquake to validate results. Other countries may not have comparable datasets for performing similar simulation studies. This data is not strictly necessary to adapt and use the proposed model for future events, but appropriate caution should be placed on results from areas where the model has not been validated. In a similar vein, the results from this study assume that survey clusters are randomly spread throughout the affected area. The decision to use randomly sampled locations was based off the assumption that systematically designing a survey sampling scheme is beyond the scope of the first few weeks of disaster response activities. In reality, there is a strong chance that rapid damage assessments might be preferentially sampled in areas with easy access, places where qualified engineers are already present, or locations with prior knowledge of severe impacts. Model performance in any of these scenarios may be worse than presented here. Adapting the model to handle preferential sampling is a clear next step for this research.

Despite these limitations, this study lays the foundation for several promising directions of future research. Most notably, this study focused on a binomial response of a single damage state. A more complex version of the model could incorporate joint likelihoods on several ordinal damage states to provide more comprehensive impact estimates. This improvement would require more detailed rapid damage assessments, but is likely to be useful even along a 'no damage—partial damage—complete damage' spectrum. The existing model could also be extended to fatality modeling. Fatalities are generally derived as the percentage of population in collapsed buildings which is simply further subset of complete damage states (Kircher et al., 2006, Robinson et al., 2018). A procedure for selecting fatality rates at the grid level would need to be developed, but otherwise the approach remains the same. Finally, model extensions that directly incorporate data on housing types from censuses or other sources could be considered. Including percentages of different housing types as an areal covariate similar to Utazi et al. (2018a) is a potential first step in this direction.

4.7 Acknowledgments

This research was performed under a National Science Foundation Graduate Research Fellowship, grant no. DGE-1450079.

4.8 Works Cited

Bhatt, S., Weiss, D. J., Cameron, E., Bisanzio, D., Mappin, B., Dalrymple, U., Battle, K. E., Moyes, C. L., Henry, A., Eckhoff, P. A., Wenger, E. A., Brit, O., Penny, M. A., Smith, T. A., Bennett, A., Yukich, J., Eisele, T. P., Griffin, J. T., Fergus, C. A., Lynch, M., Lindgren, F., Cohen, J. M., Murray, C. L. J., Smith, D. L., Hay, S. I., Cibulskis, R. E., and Gething, P. W. The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. *Nature*, 526(7572):207–211, Oct. 2015. doi: 10.1038/nature15535.

Bosco, C., Alegana, V., Bird, T., Pezzulo, C., Bengtsson, L., Sorichetta, A., Steele, J., Hornby, G., Ruktanonchai, C., Ruktanonchai, N., Wetter, E., and Tatem, A. J. Exploring the high-resolution mapping of gender-disaggregated development indicators. *Journal of The Royal Society Interface*, 14(129):20160825, Apr. 2017. ISSN 1742-5689, 1742-5662. doi: 10.1098/rsif.2016.0825.

Chaulagain, H., Rodrigues, H., Silva, V., Spacone, E., and Varum, H. Seismic risk assessment and hazard mapping in Nepal. *Natural Hazards*, 78(1):583–602, Aug. 2015. ISSN 1573-0840. doi: 10.1007/s11069-015-1734-6.

Chaulagain, H., Rodrigues, H., Silva, V., Spacone, E., and Varum, H. Earthquake loss estimation for the Kathmandu Valley. *Bulletin of Earthquake Engineering*, 14(1):59–88, Jan. 2016. ISSN 1570-761X, 1573-1456. doi: 10.1007/s10518-015-9811-5.

Comfort, L. K., Ko, K., and Zagorecki, A. Coordination in rapidly evolving disaster response systems: the role of information. In Terano, T., Kita, H., Kaneda, T., Arai, K., and Deguchi, H., editors, *Agent-Based Simulation: From Modeling Methodologies to Real-World Applications*, volume 1, pages 208–219. Springer-Verlag, Tokyo, 2005. ISBN 978-4-431-26592-4. doi: 10.1007/4-431-26925-8_19.

Erdik, M., Sestyan, K., Demircioglu, M. B., Hancilar, U., and Zulfikar, C. Rapid Earthquake Loss Assessment After Damaging Earthquakes. volume 17 of *Geotechnical, Geological, and Earthquake Engineering*, pages 523–547. Springer, 2011. doi: 10.1007/978-90-481-9544-2_21.

FEMA. Multi-hazard Loss Estimation Methodology: Earthquake Model. Technical report, Department of Homeland Security, Federal Emergency Management Agency, Washington, D.C., 2013.

Fuglstad, G.-A., Simpson, D., Lindgren, F., and Rue, H. Constructing Priors that Penalize the Complexity of Gaussian Random Fields. *Journal of the American Statistical Association*, 114 (525):445–452, Jan. 2019. ISSN 0162-1459. doi: 10.1080/01621459.2017.1415907.

Gautam, D. and Chaulagain, H. Structural performance and associated lessons to be learned from world earthquakes in Nepal after 25 April 2015 (MW 7.8) Gorkha earthquake. *Engineering Failure Analysis*, 68:222–243, Oct. 2016. ISSN 1350-6307. doi: 10.1016/j.engfailanal.2016.06.002.

Gautam, D., Prajapati, J., Paterno, K. V., Bhetwal, K. K., and Neupane, P. Disaster resilient vernacular housing technology in Nepal. *Geoenvironmental Disasters*, 3(1):1, Feb. 2016. ISSN

2197-8670. doi: 10.1186/s40677-016-0036-y.

Goodchild, M. F. and Glennon, J. A. Crowdsourcing geographic information for disaster response: a research frontier. *International Journal of Digital Earth*, 3(3):231–241, Sept. 2010. ISSN 1753-8947. doi: 10.1080/17538941003759255.

Guragain, R. *Development of Earthquake Risk Assessment System for Nepal*. PhD thesis, University of Tokyo, 2015.

Jaiswal, K. S., Wald, D. J., Earle, P. S., Porter, K. A., and Hearne, M. Earthquake casualty models within the USGS Prompt Assessment of Global Earthquakes for Response (PAGER) System. In *Human Casualties in Earthquakes*, volume 29, pages 83–94. Springer Netherlands, 2011.

Kircher, C. A., Whitman, R. V., and Holmes, W. T. HAZUS Earthquake Loss Estimation Methods. *Natural Hazards Review*, 7(2):45–59, May 2006. doi: 10.1061/(ASCE)1527-6988(2006)7:2(45).

Lallemant, D. and Kiremidjian, A. A Beta Distribution Model for Characterizing Earthquake Damage State Distribution. *Earthquake Spectra*, 31(3):1337–1352, Aug. 2015. ISSN 8755-2930. doi: 10.1193/012413EQS013M.

Lallemant, D., Soden, R., Rubinyi, S., Loos, S., Barns, K., and Bhattacharjee, G. Post-Disaster Damage Assessments as Catalysts for Recovery: A Look at Assessments Conducted in the Wake of the 2015 Gorkha, Nepal, Earthquake. *Earthquake Spectra*, 33(S1):S435–S451, Sept. 2017. ISSN 8755-2930. doi: 10.1193/120316EQS222M.

Lindgren, F. and Rue, H. Bayesian Spatial Modelling with R-INLA. *Journal of Statistical Software*, 63(1):1–25, Feb. 2015. ISSN 1548-7660. doi: 10.18637/jss.v063.i19.

Lindgren, F., Rue, H., and Lindstrm, J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498, 2011. doi: 10.1111/j.1467-9868.2011.00777.x.

Muzzini, E. and Aparicio, G. *Urban growth and spatial transition in Nepal: An initial assessment*. The World Bank, Washington, D.C., 2013.

of Nepal, G. Nepal Post Disaster Needs Assessment Report. Technical report, National Planning Commission, Kathmandu, Nepal, 2015.

Porter, K. A Beginners Guide to Fragility, Vulnerability, and Risk. In Beer, M., Kougioumtzoglou, I. A., Patelli, E., and Au, I. S.-K., editors, *Encyclopedia of Earthquake Engineering*, pages 1–29. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014. ISBN

978-3-642-36197-5. doi: 10.1007/978-3-642-36197-5_256-1.

Robinson, T. R., Rosser, N. J., Densmore, A. L., Oven, K. J., Shrestha, S. N., and Guragain, R. Use of scenario ensembles for deriving seismic risk. *Proceedings of the National Academy of Sciences*, 115(41):E9532–E9541, Oct. 2018. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1807433115.

Rue, H., Martino, S., and Chopin, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2):319–392, 2009. ISSN 1467-9868. doi: 10.1111/j.1467-9868.2008.00700.x.

Simpson, D., Rue, H., Riebler, A., Martins, T. G., and Srbye, S. H. Penalising Model Component Complexity: A Principled, Practical Approach to Constructing Priors. *Statistical Science*, 32 (1):1–28, Feb. 2017. ISSN 0883-4237, 2168-8745. doi: 10.1214/16-STS576.

Tatem, D. A., Gething, D. P., Pezzulo, D. C., Weiss, D. D., and Bhatt, D. S. Development of High-Resolution Gridded Poverty Surfaces. page 49, 2014.

Utazi, C., Thorley, J., Alegana, V., Ferrari, M., Nilsen, K., Takahashi, S., Metcalf, C., Lessler, J., and Tatem, A. A spatial regression model for the disaggregation of areal unit based data to high-resolution grids with application to vaccination coverage mapping. *Statistical Methods in Medical Research*, Sept. 2018a. doi: 10.1177/0962280218797362.

Utazi, C. E., Thorley, J., Alegana, V. A., Ferrari, M. J., Takahashi, S., Metcalf, C. J. E., Lessler, J., and Tatem, A. J. High resolution age-structured mapping of childhood vaccination coverage in low and middle income countries. *Vaccine*, 36(12):1583–1591, 2018b. doi: 10.1016/j.vaccine.2018.02.020.

Whitman, R. V., Anagnos, T., Kircher, C. A., Lagorio, H. J., Lawson, R. S., and Schneider, P. Development of a National Earthquake Loss Estimation Methodology. *Earthquake Spectra*, 13 (4):643–661, Nov. 1997. ISSN 8755-2930. doi: 10.1193/1.1585973.

WorldPop. WorldPop, 2018. URL https://www.worldpop.org/.

Xie, S., Duan, J., Liu, S., Dai, Q., Liu, W., Ma, Y., Guo, R., and Ma, C. Crowdsourcing Rapid Assessment of Collapsed Buildings Early after the Earthquake Based on Aerial Remote Sensing Image: A Case Study of Yushu Earthquake. *Remote Sensing*, 8(9):759, 2016. ISSN 2072-4292. doi: 10.3390/rs8090759.

Yeh, C.-H., Loh, C.-H., and Tsai, K.-C. Overview of Taiwan Earthquake Loss Estimation System. *Natural Hazards*, 37(1):23–37, Feb. 2006. ISSN 1573-0840. doi: 10.1007/s11069-005-4654-z.

4.9 Appendix

4.9.1 Damage Assessment Data

The damage assessment data used in this study comes from Nepal's Household Registration for Housing Reconstruction Program. In an effort spearheaded by Kathmandu Living Labs and Nepal's National Planning Commission, door to door structural damage assessments were collected at every residential household across the eleven most affected rural districts. The observed percentage of structures assessed at damage grade five are shown below in Figure 4.6. An open access version of this dataset (and the survey questions) is available at: https://eq2015.npc.gov.np/

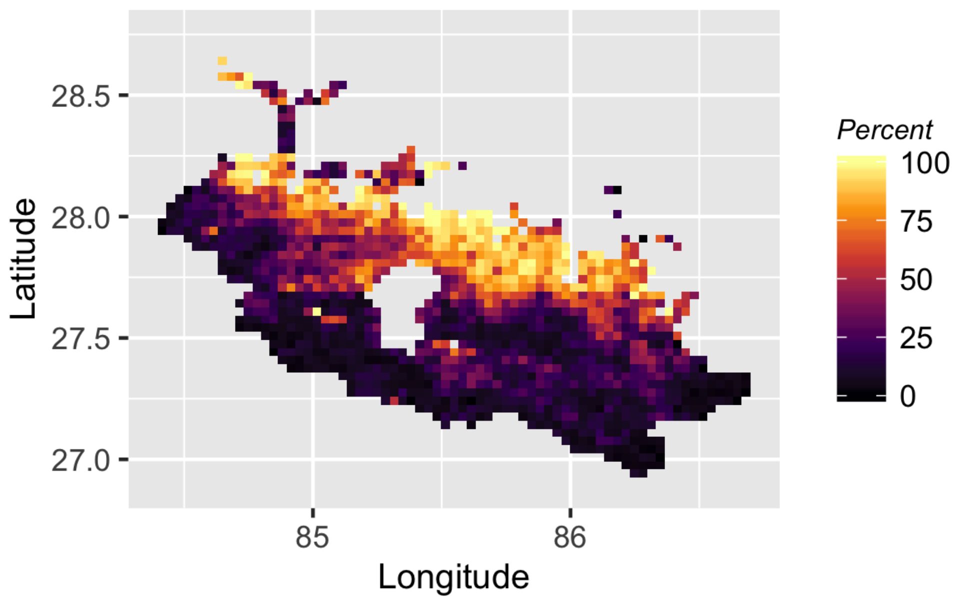

Figure 4.5: Ground truth observations for percentage of households with complete damage states in each prediction grid cell.

4.9.2 Mesh Construction

The SPDE approach requires constructing a triangulated mesh to model the spatial random field. The triangle knots serve as integration points and the values of any point lying within a triangle are interpolated from the knot estimates. Thus, a finer mesh produces more accurate predictions but increases computational time. The maximum triangle edge length for this study was set at .05 degrees or approximately 2% of the prediction grid region, similar to discretizations in previous work Utazi et al. (2018a).

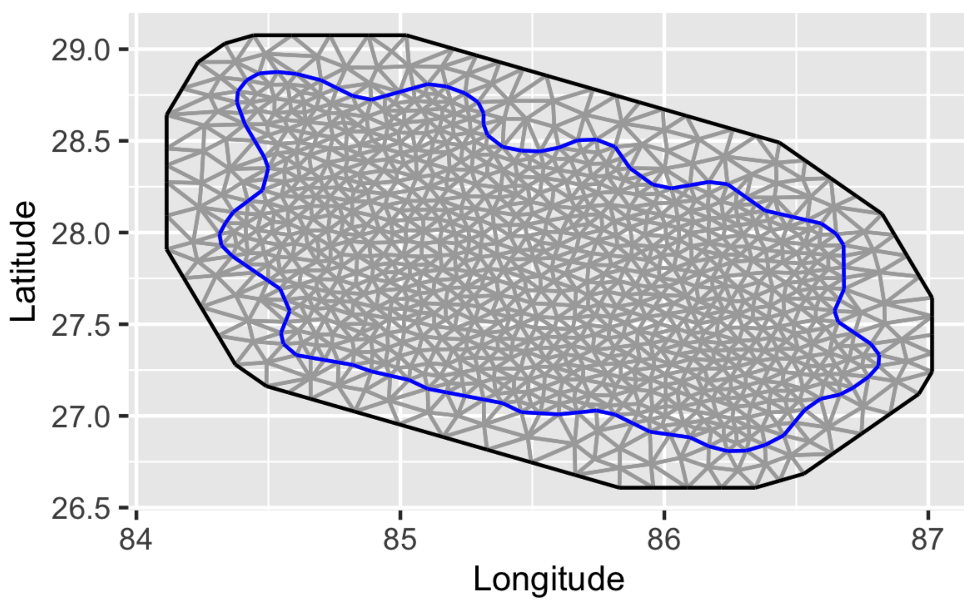

Figure 4.6: Triangular mesh used in the estimation of the spatial random field.

4.9.3 Model Parameter Estimates

Table 4.1: Median posterior parameter estimates across 50 simulations for each number of survey clusters.

Parameter	Mean	Std. Dev	95% Interval
		100 Clusters	
Intercept	-16.09	1.80	(-21.86, -11.013)
MMI	1.25	0.16	(1.00, 1.46)
log(Elevation)	0.40	0.13	(0.02, 1.69)
log(VIIRS NTL)	-0.47	0.09	(-0.74, -0.23)
log(Distance to Roads)	0.09	0.05	(-0.02, 0.20)
Range	1.45	0.06	(1.18, 1.77)
Std. Dev	2.60	0.15	(2.13, 3.15)
		500 Clusters	
Intercept	-12.74	0.57	(-13.89, -11.65)
MMI	1.37	0.04	(1.29, 1.44)
log(Elevation)	-0.12	0.03	(-0.17, -0.08)
log(VIIRS NTL)	-0.61	0.02	(-0.65, -0.56)
log(Distance to Roads)	0.26	0.01	(0.25, 0.29)
Range	1.27	0.04	(1.08, 1.50)
Std. Dev	1.20	0.08	(1.08, 1.37)
		1000 Clusters	
Intercept	-14.11823	0.43	(-14.93, -13.30)
MMI	1.86	0.02	(1.81, 1.90)
log(Elevation)	-0.15	0.02	(-0.18, -0.12)
log(VIIRS NTL)	-0.44	0.01	(-0.46, -0.42)
log(Distance to Roads)	0.38	0.01	(0.37, 0.40)
Range	1.18	0.07	(1.17, 1.33)
Std. Dev	1.03	0.07	(0.93, 1.20)

5: CONCLUSION

5.1 Review of Contributions

Using a case study of the 2015 Gorkha, Nepal Earthquake, this dissertation has highlighted three different ways in which post-disaster data can be used to inform models of earthquake vulnerabilities, reconstruction processes, and impact assessments. Each chapter used different sources of household-level data to address issues of scale and uncertainty in different portions of the disaster cycle. Although each of the chapters relies on datasets currently unique to Nepal, similar types of data are likely to be collected in future earthquakes. All model frameworks are described sufficiently generally to be adaptable to other contexts. Thus, this dissertation offers a tangible example of how post-earthquake datasets can and should be extended beyond monitoring and reporting efforts. Integrating new types of data into modeling frameworks, especially household-level surveys, provides insights that are difficult to glean from summary statistics alone. The primary insights and contributions of each chapter are summarized below.

5.1.1 Addressing scale issues in social vulnerability indices

Chapter two provided an empirical analysis of scale issue in social vulnerability indices, specifically focusing on the problems associated with using aggregated measures of household characteristics with inductive methods. This practice is pervasive among sub-national index studies due to the relative lack of micro-data sets upon which to derive actual household characteristics. Without household level data, analyses often assume that aggregated versions of the same characteristics are reasonable representation of the same theoretical components. For example, household income is assumed important and and ‘median household income’ is used as an input variable. This analysis showed that the vulnerability ranks produced using these two variables might be quite different, even if the components derived from PCA or other dimension reduction procedures are qualitatively similar. The 'average' household derived from an a mean or median may not correspond to a type of household that actually exists with any meaningful frequency. Using inductive techniques to maximizing the variance between these averaged characteristics creates components that are difficult to confidently interpret.

These results contribute to larger discussions over construct equivalence and scale sensitivities in social vulnerability indices. The notion that vulnerabilities are produced by multi-scalar processes is fundamental to several theoretical frameworks, but it remains unclear precisely how these concepts should be captured with indicators. Several existing studies have recommended using local qualitative case studies as a way to validate sub-national indices (Schmidtlein et al., 2008, Fekete, 2009, Fekete et al., 2010). This methods used in this study offer a parallel quantitative approach based in post-disaster micro-data. This study lays the foundation for further research related to measurement issues between aggregated and disaggregated scales—an important pursuit as micro-data sets become available with increased regularity.

5.1.2 Developing a modeling framework for household reconstruction behavior

Chapter three developed a new modeling approach for analyzing household level reconstruction behavior. This approach used an item-response theory framework to estimate the probability of a household taking reconstructive action from routinely collected post-disaster survey data—a novel contribution in the disaster risk reduction space. While it is common to track reconstruction progress through aid dispersal statistics or localized case studies, no existing research has attempted to generalize trends at the scale of an entire earthquake event. In owner-driven reconstruction systems, this type of research is especially important for understanding the factors that contribute to increased reconstruction probability.

The analysis in chapter three found several notable trends in an analysis of 5,913 household-level surveys collected approximately one year after reconstruction activities started in Nepal. Economic aid in the form of the first round of reconstruction funding or similar I/NGO funding had a relatively small impact on the probability of a household starting the reconstruction process (about 8% on average). Receipt of an engineering consultation was considerably more impactful for increasing reconstruction probabilities. Household willingness to commit additional resources (financial, material, or otherwise) also had a large effect on reconstruction probabilities. Households with high levels of additional resource contribution also benefited relatively more from external aid, suggesting that equal aid distribution does not necessarily correspond with equal reconstruction outcomes. These specific findings contribute to discussions of equitable earthquake recovery and highlight the need for further research that seeks to understand the barriers for reconstruction in resource-poor households. More broadly however, these types of insights also communicate the potential for using sophisticated model frameworks to evaluate reconstruction progress in much greater detail than is currently the norm.

5.1.3 Improving modeled damage estimates with geolocated cluster data

Chapter four examined an alternative damage modeling framework that relies on geolocated rapid damage assessments rather than fragility curves to estimate damage levels. Although rapid damage models are often accurate at coarse scales, they lack the ability to incorporate ground truthed information into estimation procedures. For the Gorkha Earthquake, this meant leaving out information from over 60,0000 rapid damage assessments performed within several weeks of the event. The modeling framework in this chapter proposes a spatial interpolation framework for estimating the percentage of completely damaged structures on a high resolution grid from this type of sparse survey cluster data. The model uses the INLA-SPDE approach, a relatively new computational method designed for spatial modeling (Lindgren and Rue, 2015).

The results from this study show meaningful improvements over existing damage models, reducing prediction errors by over 25% while simultaneously increasing spatial resolution. The proposed framework is able to capture the broad damage trends with approximately 5,500 rapid damage assessments—less than 10% of the assessments collected in a few weeks after the Gorkha earthquake. With tens of thousands of damage assessments, model predictions are able to capture fine-scale spatial heterogeneities in damage patterns. These findings represent a promising new direction for rapid damage estimation by leveraging damage assessments that are already being collected after major earthquakes. The proposed methods can be implemented within the timeline for delivering a PDNA and offer damage estimates based in actual field observations. Translating these model results into more direct loss metrics and aid estimates remains an area of future work, but this chapter lays the necessary groundwork for such developments. This work also contributes to broader active research efforts focused on synthesizing the new wave of rapidly available post-disaster impact data (Goodchild and Glennon, 2010, Xie et al., 2016, Lallemant et al., 2017, Li et al., 2019).

5.2 Limitations and Opportunities

This dissertation uses the 2015 Gorkha earthquake as a case study for how various post-disaster datasets could improve different aspects of disaster modeling. However, relying exclusively on datasets from one event and collected for other purposes brings certain limitations to each chapter. Some of the most important points are summarized below:

- Chapter two relied on eleven variables extracted from the socio-demographic data collected in the household reconstruction surveys. This is fewer variables than would normally be used in similar indices based on census data. Although most of the same dimensions are represented, the paper is largely positioned as a methodological critique rather than an actual social vulnerability index due to the limited variable set. Fortunately, including additional variables in future work (if available) would not require any adjustments. Synthesizing micro-data sets with census data across common variables is also likely to be a useful topic of future research.

- Chapter three used three months of survey data collected approximately one year after reconstruction activities started in May of 2016. As a result, the model evaluated a single snapshot of reconstruction progress amongst a three to five year reconstruction timeline. Collection of CFP surveys continued further into the reconstruction process, but changes in survey questions and Nepal's administrative restructuring prevented the incorporation of more recent surveys into the analysis. This could be improved in future events by maintaining a standard set of questions for reconstruction surveys collected across several years. Reconstruction continues to be one of the most under-researched components of earthquake DRR due to a lack of consistent longitudinal datasets. Facilitating more partnerships and data sharing between academic researchers and stakeholders on the ground is a clear step towards improving general understanding of reconstruction processes.

- Chapter four only considered completely damaged buildings in its damage estimates rather than more traditional ordinal damage states. This decision was made on the basis that fully damaged and collapsed buildings are straightforward to rapidly identify and in need of the most immediate assistance. A more complex statistical model involving joint likelihoods for multiple damage states is likely feasible, but would requires significant further development, testing, and validation. Determining how to best link geospatial covariates to ordinal damages states is a particularly challenging area of future research. Working out methods for integrating housing census data into predictions is another direction worth pursuing.

- Chapters three and four used specific domain knowledge about the study region during the model development processes. For chapter three, this included research on the factors affecting reconstruction rates in rural environments and specific details of Nepal's reconstruction framework. For chapter four, this involved selecting gridded covariates that correlated strongly with different housing types. As a result, neither of these models can be carbon copied into an alternate context without modification. The model frameworks are described in sufficient detail in both studies to be adapted to different settings, but model performance may vary. Applying these models to other events is an important future step for validating the generalizability of the proposed frameworks.

- This dissertation relied on survey data that was collected for reporting and monitoring purposes. Reformatting these datasets for modeling applications required a significant amount of work. For example, the CFP surveys used in Chapter three required a combination of fuzzy-matching and manual joins on village development committee names translated from Nepali before they could be linked to other datasets. Improving the machine readability of datasets would increase the uptake of new datasets and reduce the potential for errors in the data preparation phase of statistical analyses. Standardizing administrative coding schemes and reporting of survey non-response alone would go a long ways in improving the usability of various datasets. Providing open access to datasets is important, but reducing the corresponding barriers to entry is equally important.

5.3 Towards an improved earthquake disaster modeling ecosystem

The ecosystem for post-event data collection and distribution has drastically improved in recent earthquakes and is primed to continue expanding. Access to crowd-sourcing platforms and and remotely sensed imagery has never been better and open-data sharing is steadily becoming a standard practice among NGOs and government agencies. This dissertation focused on providing diverse examples of how earthquake disaster models can be improved with novel varieties of post-disaster data rather than narrowing in on any one specific area. It is hoped that these studies motivate future research in similar directions, whether further developing these contributions or innovating new approaches altogether. Taking full advantage of the disaster data explosion will require new techniques, collaborations, and a desire to dive into the trenches of messy datasets. For those willing however, the current surge of new data streams offers an unparalleled opportunity to understand the dynamisms of disaster processes at high resolution and across multiple scales. These insights are critical for mitigating hazards, managing risks, and reducing vulnerabilities for those living amongst the world's most active fault zones.

6: WORKS CITED

Alegana, V. A., Wright, J., Pezzulo, C., Tatem, A. J., and Atkinson, P. M. Treatment-seeking behaviour in low- and middle-income countries estimated using a Bayesian model. *BMC Medical Research Methodology*, 17(1), Dec. 2017. doi: 10.1186/s12874-017-0346-0.

Alegana, V. A., Maina, J., Ouma, P. O., Macharia, P. M., Wright, J., Atkinson, P. M., Okiro, E. A., Snow, R. W., and Tatem, A. J. National and sub-national variation in patterns of febrile case management in sub-Saharan Africa. *Nature Communications*, 9(1), Dec. 2018. doi: 10.1038/s41467-018-07536-9.

Amin, S. and Goldstein, M. P., editors. *Data against natural disasters: establishing effective systems for relief, recovery, and reconstruction*. World Bank, Washington DC, 2008. ISBN 978-0-8213-7452-8 978-0-8213-7453-5. OCLC: ocn185123397.

Beccari, B. A Comparative Analysis of Disaster Risk, Vulnerability and Resilience Composite Indicators. *PLoS Currents*, 2016. doi: 10.1371/currents.dis.453df025e34b682e9737f95070f9b970.

Bownas, R. and Bishokarma, R. Access after the earthquake: the micro politics of recovery and reconstruction in Sindhupalchok District, Nepal, with particular reference to caste. *Contemporary South Asia*, 0(0):1–17, Dec. 2018. doi: 10.1080/09584935.2018.1559278.

Comfort, L. K. and Joshi, J. Scalability and Sustainability in Uncertain Environments: Transition to Recovery from the 2015 Gorkha, Nepal, Earthquakes. *Earthquake Spectra*, 33(S1): S385–S401, Dec. 2017. doi: 10.1193/113016EQS217M.

Cutter, S. L., Boruff, B. J., and Shirley, W. L. Social Vulnerability to Environmental Hazards. *Social Science Quarterly*, 84(2):242–261, June 2003. doi: 10.1111/1540-6237.8402002.

Daly, P., Ninglekhu, S., Hollenbach, P., Duyne Barenstein, J., and Nguyen, D. Situating local stakeholders within national disaster governance structures: rebuilding urban neighbourhoods following the 2015 Nepal earthquake. *Environment and Urbanization*, 29(2):403–424, Oct. 2017. doi: 10.1177/0956247817721403.

Dangal, R. Country profile: Nepal. Technical report, Asian Disaster Reduction Center, 2011.

Erdik, M., Sestyan, K., Demircioglu, M. B., Hancilar, U., and Zulfikar, C. Rapid Earthquake Loss Assessment After Damaging Earthquakes. volume 17 of *Geotechnical, Geological, and Earthquake Engineering*, pages 523–547. Springer, 2011. doi: 10.1007/978-90-481-9544-2_21.

Fekete, A. Validation of a social vulnerability index in context to river-floods in Germany. *Natural Hazards and Earth System Sciences*, 9(2):393–403, Mar. 2009. doi:

10.5194/nhess-9-393-2009.

Fekete, A., Damm, M., and Birkmann, J. Scales as a challenge for vulnerability assessment. *Natural Hazards*, 55(3):729–747, 2010. doi: 10.1007/s11069-009-9445-5.

Goda, K., Kyota, T., Pokhrel, R. M., Chiaro, G., Katagiri, T., Sharma, K., and Wilkinson, S. The 2015 Gorkha Nepal earthquake: insights from earthquake damage survey. *Frontiers in Built Environment*, 1(8):1–15, 2015.

Goodchild, M. F. and Glennon, J. A. Crowdsourcing geographic information for disaster response: a research frontier. *International Journal of Digital Earth*, 3(3):231–241, Sept. 2010. doi: 10.1080/17538941003759255.

Government of Nepal. Nepal Post Disaster Needs Assessment Report. Technical report, National Planning Commission, Kathmandu, Nepal, 2015.

Government of Nepal. Nepal Earthquake 2015: Post Disaster Recovery Framework 2016-2020. Technical report, Government of Nepal, Kathmandu, Nepal, Apr. 2016.

Hall, M., Lee, A., Cartwright, C., Marahatta, S., Karki, J., and Simkhada, P. The 2015 Nepal earthquake disaster: lessons learned one year on. *Public Health*, 145:39–44, Apr. 2017. doi: 10.1016/j.puhe.2016.12.031.

He, L., Aitchison, J. C., Hussey, K., Wei, Y., and Lo, A. Accumulation of vulnerabilities in the aftermath of the 2015 Nepal earthquake: Household displacement, livelihood changes and recovery challenges. *International Journal of Disaster Risk Reduction*, 31:68–75, Oct. 2018. doi: 10.1016/j.ijdrr.2018.04.017.

Jaiswal, K. S., Wald, D. J., Earle, P. S., Porter, K. A., and Hearne, M. Earthquake casualty models within the USGS Prompt Assessment of Global Earthquakes for Response (PAGER) System. In *Human Casualties in Earthquakes*, volume 29, pages 83–94. Springer Netherlands, 2011.

Joshi, G. R. and Joshi, N. B. Economic Loss from Earthquake in Nepal and Strategies for Recovery and Resilience Building. In *Living Under the Threat of Earthquakes*, Springer Natural Hazards, pages 195–209. Springer, Cham, 2018. ISBN 978-3-319-68043-9 978-3-319-68044-6. doi: 10.1007/978-3-319-68044-6_13.

Kessler, C. and Hendrix, C. The Humanitarian eXchange Language: Coordinating Disaster Response with Semantic Web Technologies. *Semantic Web*, 1:1–5, 2009.

Khazai, B., Kunz-Plapp, T., Bscher, C., and Wegner, A. VuWiki: An Ontology-Based Semantic Wiki for Vulnerability Assessments. *International Journal of Disaster Risk Science*, 5(1): 55–73, Mar. 2014. doi: 10.1007/s13753-014-0010-9.

Lallemant, D., Soden, R., Rubinyi, S., Loos, S., Barns, K., and Bhattacharjee, G. Post-Disaster Damage Assessments as Catalysts for Recovery: A Look at Assessments Conducted in the Wake of the 2015 Gorkha, Nepal, Earthquake. *Earthquake Spectra*, 33(S1):S435–S451, Sept. 2017. doi: 10.1193/120316EQS222M.

Li, G., Zhao, J., Murray, V., Song, C., and Zhang, L. Gap analysis on open data interconnectivity for disaster risk research. *Geo-spatial Information Science*, 22(1):45–58, Jan. 2019. doi: 10.1080/10095020.2018.1560056.

Lindgren, F. and Rue, H. Bayesian Spatial Modelling with R-INLA. *Journal of Statistical Software*, 63(1):1–25, Feb. 2015. ISSN 1548-7660. doi: 10.18637/jss.v063.i19.

Lindgren, F., Rue, H., and Lindstrm, J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498, 2011. doi: 10.1111/j.1467-9868.2011.00777.x.

Mishra, A., Ghate, R., Maharjan, A., Gurung, J., Pathak, G., and Upraity, A. N. Building ex ante resilience of disaster-exposed mountain communities: Drawing insights from the Nepal earthquake recovery. *International Journal of Disaster Risk Reduction*, 22:167–178, June 2017. doi: 10.1016/j.ijdrr.2017.03.008.

Parajuli, R. R. and Kyono, J. Ground motion characteristics of the 2015 Gorkha earthquake, survey of damage to stone masonry structures and structural field tests. *Frontiers in Built Environment*, 1(23):1–12, 2015.

Rai, D. C., Singha, V., S, B. R., and Sager, S. L. Reconnaissance of the effects of the M7.8 Gorkha (Nepal) earthquake of April 25, 2015. *Geomatics, Natural Hazards and Risk*, 7(1): 1–17, 2016.

Rue, H., Martino, S., and Chopin, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2):319–392, 2009. doi: 10.1111/j.1467-9868.2008.00700.x.

Schmidtlein, M. C., Deutsch, R. C., Piegorsch, W. W., and Cutter, S. L. A sensitivity analysis of the social vulnerability index. *Risk Analysis*, 28(4):1099–1114, 2008. doi: 10.1111/j.1539-6924.2008.01072.x.

Tate, E. Social vulnerability indices: a comparative assessment using uncertainty and sensitivity analysis. *Natural Hazards*, 63:325–347, 2012. doi: 10.1007/s11069-012-0152-2.

Tate, E. Uncertainty Analysis for a Social Vulnerability Index. *Annals of the Association of*

American Geographers, 103(3):526–543, May 2013. doi: 10.1080/00045608.2012.700616.

Thomas, D. S. K., Phillips, B. D., Lovekamp, W. E., Fothergill, A., Phillips, B. D., Lovekamp, W. E., and Fothergill, A. *Social Vulnerability to Disasters*. CRC Press, May 2013. ISBN 978-1-4665-1638-0. doi: 10.1201/b14854.

UNDRR. Sendai Framework for Disaster Risk Reduction, 2015. URL https://www.unisdr.org/we/coordinate/sendai-framework.

Wisner, B., Blaikie, P., Cannon, T., and Davis, I. *At Risk: Natural hazards, people's vulnerability and disasters*. Routeldge, 2nd edition, 2004.

World Bank. Guide to Developing Disaster Recovery Frameworks: Sendai Conference Version. Technical report, World Bank, Washington, DC., Mar. 2015.

Xie, S., Duan, J., Liu, S., Dai, Q., Liu, W., Ma, Y., Guo, R., and Ma, C. Crowdsourcing Rapid Assessment of Collapsed Buildings Early after the Earthquake Based on Aerial Remote Sensing Image: A Case Study of Yushu Earthquake. *Remote Sensing*, 8(9):759, 2016. doi: 10.3390/rs8090759.

Zhang, H., Qing, X., Huang, M., and Li, G. A Correlation Analysis Model for Multidisciplinary Data in Disaster Research. *Data Science Journal*, 14(0):19, May 2015. doi: 10.5334/dsj-2015-019.