Modeling Hurricane Evacuation Behavior using a Dynamic Discrete Choice Framework

Tarun Rambha*[1,2], Linda K. Nozick[3], and Rachel Davidson[4]

[1]Department of Civil Engineering, Indian Institute of Science
[2]Center for infrastructure, Sustainable Transportation and Urban Planning, Indian Institute of Science
[3]School of Civil and Environmental Engineering, Cornell University
[4]Civil and Environmental Engineering, University of Delaware

Abstract

Predicting evacuation-related choices of households during a hurricane is of paramount importance to any emergency management system. Central to this problem is the identification of socio-demographic factors and hurricane characteristics that influence an individual's decision to stay or evacuate. However, decision makers in such conditions do not make a single choice but constantly evaluate current and anticipated conditions before opting to stay or evacuate. We model this behavior using a finite-horizon dynamic discrete choice framework in which households may choose to evacuate or wait in time periods prior to a hurricane's landfall. In each period, an individual's utility depends not only on his/her current choices and the present values of the influential variables, but also involves discounted expected utilities from future choices should one decide to postpone their decision to evacuate. Assuming generalized extreme value (GEV) errors, a nested algorithm involving a dynamic program and a maximum likelihood method is used to estimate model parameters. Panel data on households affected by Hurricane Gustav, which made landfall in Louisiana on 1 September 2008, was fused with the National Hurricane Center's forecasts on the trajectory and intensity for the case study in the paper.

Keywords: Hurricanes; evacuation, demand estimation, dynamic discrete choice, maximum-likelihood

1 Introduction

Almost every year, Atlantic hurricanes that strike the US coast leave a trail of destruction and result in the loss of hundreds of lives. Disasters such as Harvey and Katrina stand as examples of the catastrophic nature of these events and have been estimated to cost between $108 and $198 billion (Knabb et al., 2011; Hicks and Burton, 2017). In situations like these, preparedness is vital to minimize casualties and economic and infrastructural losses. Fortunately, unlike other natural events, hurricanes can be tracked several days in advance, allowing individuals to vacate potentially hazardous regions and governments to deploy disaster management strategies. However, mass evacuations are easier said than done.

In the last two decades, major hurricanes such as Floyd, Rita, Gustav (and the more recent Irma and Michael) have each prompted the evacuation of nearly 2 to 3 million residents from various coastal states (Knabb et al., 2006; Gottumukkala et al., 2011). At an individual or household level, the decision to evacuate or stay is hard because of the uncertainty in the trajectory and the intensity of a hurricane. Leaving early may not be the best thing to do if a hurricane changes its path or if it beings to abate. On the other hand, waiting longer or choosing to stay is risky because hurricanes are accompanied by strong winds and flooding, which can result in major disruptions to power, water, roadways, and other supply-chains.

At an aggregate level, evacuations are extremely difficult to plan, execute, and monitor due to the scale and the interplay between complex interconnected networks. During times of disasters, roadways that are typically designed to handle regular day-to-day traffic witness massive traffic jams that may sometimes take 10–20 hours to clear (Lindell

*✉ tarunrambha@iisc.ac.in

et al., 2005; Soika, 2006). In addition, providing real-time updates to individuals regarding a hurricane or the network conditions becomes increasingly difficult. Such situations also make it impossible to transport emergency personnel and essential supplies such as food and fuel.

One of the major challenges in planning an evacuation is to understand who evacuates and why. Identifying the factors which influence evacuation decisions and being able to predict the overall demand can assist public agencies in issuing mandatory evacuation orders (Yi et al., 2017), contra-flow plans (Wolshon, 2001; Tuydes and Ziliaskopoulos, 2006), and in design and operations of shelter locations (Li et al., 2011). Using econometric models, several previous studies have analyzed the impacts of socio-demographic variables (such as household income, composition, education, vehicle ownership, and presence of pets, children, and elders) and hurricane-related characteristics (such as intensity, distance, storm surge, and wind speeds) on the propensity to evacuate. However, emergency response to hurricanes is a time-bound problem, and the question of when individuals evacuate is equally important but has received relatively little attention.

During the time periods leading up to a hurricane's landfall, individuals can choose to wait or evacuate based on their household characteristics and current and predicted hurricane and traffic information. Thus, when addressing the evacuation demand estimation problem using a discrete choice framework, the utility functions of decision makers must feature future estimates of covariates in addition to the present conditions. Our current paper is geared towards filling this gap, and using a finite-horizon dynamic choice model, we represent the time dynamics of an individual's evacuation behavior during a hurricane. We also derive estimates of time-varying choice probabilities and marginals/sensitivities, which offer valuable insights for selecting higher-level disaster management plans.

The remainder of the paper is structured as follows. In Section 2, we review literature on demand estimation of hurricane evacuees and specifically discuss relevant econometric models that have been used in the past. Section 3 provides background on dynamic discrete choice theory. Section 4 formulates the evacuation process using this framework and details the maximum likelihood method for three models: one with prefect information and two with partial information. In Section 5, we apply this procedure to estimate model parameters from a household survey on evacuation behavior of residents affected by Hurricane Gustav, which made landfall in Louisiana on 1 September 2008. We also present results on the sensitivities of the choice probabilities and validate our models. Finally, Section 6 summarizes our findings and suggests potential extensions to the current research.

2 Literature Review

Hurricanes are tropical cyclones that originate in the Atlantic Ocean with maximum sustained winds that exceed 74 mph. They are typically classified on a scale of 1 to 5, called the Saffir–Simpson scale, with higher numbers indicating stronger storms. The National Hurricane Center (NHC) tracks all Atlantic tropical storms and provides updates every 6 hours with predictions on track cone and the probabilities with which the wind speeds and storm surge may exceed different thresholds (`http://www.nhc.noaa.gov/`). Several state and local agencies often use this information to issue advisories and optional and mandatory evacuation orders that help vulnerable populations to evacuate to safer regions. Common destination choices during hurricanes include hotels, houses of friends and family, and shelters. Detailed summaries on hurricanes and associated evacuation issues can be found in Dash and Gladwin (2007), Pel et al. (2012), Murray-Tuite and Wolshon (2013), and Rambha et al. (2019).

A key to facilitating the evacuation of large populations lies in the estimation of travel demand, a process which is commonly referred to as trip generation. Traditionally, this was carried out using two stages in which the first stage involves computing participation rates from past evacuation stats and census data. In the second stage, the cumulative demand was assumed to be distributed over time according to S-shaped patterns (FEMA, 1999). Commonly used S-shaped curves include Weibull (Lindell, 2008) and Sigmoid (Xie et al., 2010) distributions. While estimating demands using the two-stage method is simple (and perhaps applicable to other types of natural hazards such as nuclear disasters), it suffers from a major drawback. Hurricanes offer lead times that could span several days, and empirical data suggests that the time of day and issuance of mandatory orders creates peaks in the departure rates which cannot be modeled using S-shaped curves (Lindell and Prater, 2007; Fu et al., 2007; Dixit et al., 2011).

A more promising line of research involves using statistical methods to compute the odds with which an individual or a household would evacuate. Frequently used predictors for this purpose include household composition and characteristics, type of construction, the risk level of the area, evacuation orders, and storm-specific threats such as intensity and location (Baker, 1991; Peacock et al., 1997; Whitehead et al., 2000; Lindell et al., 2011). Statistical models also help in understanding the impacts of many other not-so-distinct factors. For instance, Lindell et al. (2005) and Dow and Cutter (1998) identified different sources of information and analyzed their relative importance in the context of evacuations. Likewise, Widener et al. (2013) and Sadri et al. (2017) proposed agent-based and mixed logit models respectively to gain insights into the role of social networks on information dissemination during hurricanes. Some studies also examined the effects of back-to-back hurricanes on the choices of households (Dow and Cutter, 1998; Wu et al., 2012). For example, the extent of damage caused by Hurricanes Katrina and Harvey played a significant role in triggering massive evacuations during Hurricanes Rita and Irma that followed them. From a methodological standpoint, discrete choice models such as logit (Whitehead et al., 2000; Wilmot and Mei, 2004), logit with random parameters (Hasan et al., 2011; Sarwar et al., 2016), probit (Vásquez et al., 2015), and ordered probit (Xu et al., 2016) have been used to study individuals' binary decisions to evacuate or not. Choice models were found to perform better than participation rates in predicting evacuation decisions (Wilmot and Mei, 2004; Xu et al., 2016). Transferability of econometric approaches in the context of hurricane evacuations (Hasan et al., 2011) and nonlinear disutilities to capture risk attitudes of decision makers (Dixit et al., 2012) have also been addressed in the literature.

Few papers on hurricane demand estimation focus on the distribution of evacuation decisions over time. Fu and Wilmot (2004) looked at the joint problem of evacuating and selecting the departure time using a sequential logit model. They divided time into a finite number of intervals and assumed that decision makers choose between evacuating and staying in each time period. The associated probabilities for a particular time period were estimated using a binary logit model after discarding the data on individuals who evacuated in previous periods. The sequential logit model can, however, lead to a mismatch between the predicted and the observed number of evacuees. To resolve this issue, Gudishala and Wilmot (2012) proposed a time-dependent nested logit model in which the choice of evacuating or staying at each time interval belonged to a nest representing the stay option from the previous time interval. While the nesting structure allowed the authors to capture correlations between error terms across time periods, decision makers choose between terminal alternatives in one shot, and the values of the covariates are assumed to be deterministic and are known at every level/time period. The coefficients of the utility expressions and the log sum parameters of nested models are also notoriously difficult to estimate, especially for the hurricane demand estimation problem, which may involve several time periods.

In a different approach, Sarwar et al. (2016) used indicator variables to represent multiple time intervals before landfall and simultaneously modeled evacuation and departure time choices using a binary logit model with random parameters. Dynamic discrete choice models in which individuals decide between evacuating and staying in different time periods were also suggested by Czajkowski (2011) and Serulle and Cirillo (2017). The former paper defines a risk index measure which is a single value that combines future hurricane track and intensity information. The risk measures are used at each stage to figure out if a household must evacuate. In the latter paper, households are assumed to make choices not just on the current values of the independent variables but also factor in the utilities they may receive from postponing evacuation decisions. While our research adopts a similar modeling framework, our formulation and solution techniques differ in several ways, which will be highlighted in Section 4.

Dynamic discrete choice was popularized in Rust's seminal paper on the replacement of bus engines (Rust, 1987). According to this framework, it is assumed that individuals maximize their expected discounted utilities over a finite or an infinite time horizon akin to a Markov decision process (MDP). However, the rewards and the transition probabilities of this MDP are not known, and the task of the analyst is to uncover them using panel data on the decisions and data on a set of independent variables that influence one's choices. Estimating such models can be carried out using a nested fixed-point algorithm (Rust, 1988, 1994) or a conditional choice probability estimator (Hotz and Miller, 1993). Since its conception, dynamic choice theory has had several applications in economics: employment decisions (Heckman, 1981; Miller, 1984); fertility behavior, i.e., choice of number and timing of children (Wolpin, 1984); renewing patents (Pakes, 1984); modeling demand for durable products such as video games, printers, and camcorders (Nair, 2007; Melnikov, 2013; Gowrisankaran and Rysman, 2012); identifying the effects of student employment on an individual's success in academic and labor markets (Joensen, 2009), and analyzing the timing and choice of construction in housing markets (Murphy, 2015). Several of these problems share an optimal stopping property in which a certain action terminates the decision-making process, a feature that is common to the evacuation problem as well. Dynamic discrete choice models have also been used in transportation research to study car ownership (Glerum et al., 2015; Cirillo et al., 2015; Liu and Cirillo, 2018) and route choice (Mai et al., 2015). The reader is encouraged to see Aguirregabiria and Mira (2010) and Arcidiacono and Ellickson (2011) for detailed reviews on this framework.

3 Dynamic Choice Models

3.1 Background

In a dynamic choice setting, decision makers are assumed to choose from a finite set of alternatives repeatedly over time. The time period of interest is divided into intervals $t = 1, 2, \ldots, T$. In period t, the decision maker can select an action a_t from a set of actions A_t after observing a state vector $(\mathbf{y}_t, \boldsymbol{\epsilon}_t)$. Throughout this section, we deal with a single decision maker's choices and hence avoid indexes corresponding to individuals in the notation. The vector $\mathbf{y}_t$ is assumed to contain both individual- and alternative-specific attributes and is observable to both the decision maker and the analyst. On the other hand, the $\boldsymbol{\epsilon}_t$ vector consists of one term for each action $\epsilon_t(a_t)$, which is observable only to the decision maker and captures the latent attributes that influence the choice-making process. Let X_t represent the state space in time period t, which is the set of all values that the state vector can take. We will use $\mathbf{x}_t$ to denote a generic state and $\mathbf{y}_t$ to represent a realization of the state or data. We assume that individuals receive one-step utilities that are additive separable, $u(\mathbf{x}_t, a_t) + \epsilon_t(a_t)$, for selecting an action a_t at state $(\mathbf{x}_t, \boldsymbol{\epsilon}_t)$.[1]

The objective of the dynamic choice framework is to parameterize the observable components of the one-step utilities and use a maximum likelihood estimation method to understand the effect of different attributes on the choice making process. To this end, we need the probability of selecting different actions $\mathbb{P}[a_t|\mathbf{y}_t]$ in each time period. Just as in static choice models, the choice probabilities can be expressed as

$$\mathbb{P}[a_t|\mathbf{y}_t] = \mathbb{P}[v_t(\mathbf{y}_t, a_t) + \epsilon_t(a_t) > v_t(\mathbf{y}_t, a_t') + \epsilon_t(a_t') \, \forall \, a_t' \in A_t \backslash \{a_t\}] \tag{1}$$

where $v_t(\mathbf{y}_t, a_t)$ is the *conditional value function*, which is a measure of the utility from choosing a_t in time period t and behaving "optimally" thereafter. In addition, if the error terms $\epsilon_t(a_t)$ are independent across alternatives and identically distributed according to extreme value type I distribution, we may write

$$\mathbb{P}[a_t|\mathbf{y}_t] = \frac{\exp\left(v_t(\mathbf{y}_t, a_t)\right)}{\sum_{a_t' \in A_t} \exp\left(v_t(\mathbf{y}_t, a_t')\right)} \tag{2}$$

However, unlike static choice models, $v_t(\mathbf{y}_t, a_t)$ cannot be expressed in closed form but is obtained by solving a dynamic program. In the rest of this section, we focus our attention on the steps involved in finding these conditional value functions for Markovian transition functions. Other extensions will be discussed in Section 4.

Recall that the decision maker selects an action after observing a state $(\mathbf{y}_t, \boldsymbol{\epsilon}_t)$ in time period t.[2] Suppose a sequence of actions taken by the decision maker, also called a *policy*, is represented by a vector of functions $\boldsymbol{\pi} = \big(\pi_1(.), \pi_2(.), \ldots, \pi_T(.)\big)$, where the function $\pi_t(.)$ takes an argument $(\mathbf{x}_t, \boldsymbol{\epsilon}_t)$ and returns an action in A_t. Then, the utility of a policy $\boldsymbol{\pi}$ can be expressed as

$$\sum_{t=1}^{T} \alpha^{t-1} \Big\{ u\big(\mathbf{x}_t, \pi_t(\mathbf{x}_t, \boldsymbol{\epsilon}_t)\big) + \epsilon_t\big(\pi_t(\mathbf{x}_t, \boldsymbol{\epsilon}_t)\big) \Big\} \tag{3}$$

where α represents the discount factor. Suppose that Π denotes the set of all admissible policies. Then, the objective of the decision maker is

$$\max_{\boldsymbol{\pi} \in \Pi} \mathbb{E} \sum_{t=1}^{T} \alpha^{t-1} \Big\{ u\big(\mathbf{x}_t, \pi_t(\mathbf{x}_t, \boldsymbol{\epsilon}_t)\big) + \epsilon_t\big(\pi_t(\mathbf{x}_t, \boldsymbol{\epsilon}_t)\big) \Big\} \tag{4}$$

where the expectation is taken over the distributions of $\mathbf{x}$ and $\boldsymbol{\epsilon}$. The ϵ's are assumed to be i.i.d over time periods with a probability density function $g(\boldsymbol{\epsilon}_t)$ and the observable component of the state vector $\mathbf{x}$ is assumed Markovian with a probability density $f(\mathbf{x}_{t+1}|\mathbf{x}_t, \boldsymbol{\epsilon}_t, a_t)$. However, for the sake of tractability, it is assumed that $f(\mathbf{x}_{t+1}|\mathbf{x}_t, \boldsymbol{\epsilon}_t, a_t) = f(\mathbf{x}_{t+1}|\mathbf{x}_t, a_t)$, which is commonly referred to as the *conditional independence* assumption.

The Bellman's optimality conditions for the above problem can be written, using value functions $V_t(\mathbf{x}_t, \boldsymbol{\epsilon}_t)$, as follows:

$$V_t(\mathbf{x}_t, \boldsymbol{\epsilon}_t) = \max_{a_t \in A_t} \left\{ u(\mathbf{x}_t, a_t) + \epsilon_t(a_t) + \alpha \int \left(\int V_{t+1}(\mathbf{x}_{t+1}, \boldsymbol{\epsilon}_{t+1}) g(\boldsymbol{\epsilon}_{t+1}) d\boldsymbol{\epsilon}_{t+1} \right) f(\mathbf{x}_{t+1}|\mathbf{x}_t, a_t) d\mathbf{x}_{t+1} \right\} \tag{5}$$

[1] The observable components of the one-step utilities $u(\mathbf{x}_t, a_t)$ can also be assumed to be explicitly time-dependent but estimating such a model is relatively difficult.

[2] When $\boldsymbol{\epsilon}_t$ appears alongside $\mathbf{y}_t$ it is assumed to be a realization of the vector of error terms and when it appears alongside $\mathbf{x}_t$, it is a random variable. We avoid additional syllabary to distinguish these two cases for notational simplicity.

The issue with finding the value functions $V_t(\mathbf{x}_t, \boldsymbol{\epsilon}_t)$ is that the ϵ's are continuous random variables and hence solving the dynamic program is computationally difficult. We therefore integrate out the ϵ's and define *ex ante value functions* $\overline{V}_t(\mathbf{x}_t)$ as shown below.

$$\overline{V}_t(\mathbf{x}_t) = \int V_t(\mathbf{x}_t, \boldsymbol{\epsilon}_t) g(\boldsymbol{\epsilon}_t) d\boldsymbol{\epsilon}_t \tag{6}$$

From equations (5) and (6), we may write

$$V_t(\mathbf{x}_t, \boldsymbol{\epsilon}_t) = \max_{a_t \in A_t} \left\{ u(\mathbf{x}_t, a_t) + \epsilon_t(a_t) + \alpha \int \overline{V}_{t+1}(\mathbf{x}_{t+1}) f(\mathbf{x}_{t+1}|\mathbf{x}_t, a_t) d\mathbf{x}_{t+1} \right\} \tag{7}$$

$$\int V_t(\mathbf{x}_t, \boldsymbol{\epsilon}_t) g(\boldsymbol{\epsilon}_t) d\boldsymbol{\epsilon}_t = \int \max_{a_t \in A_t} \left\{ u(\mathbf{x}_t, a_t) + \epsilon_t(a_t) + \alpha \int \overline{V}_{t+1}(\mathbf{x}_{t+1}) f(\mathbf{x}_{t+1}|\mathbf{x}_t, a_t) d\mathbf{x}_{t+1} \right\} g(\boldsymbol{\epsilon}_t) d\boldsymbol{\epsilon}_t \tag{8}$$

$$\overline{V}_t(\mathbf{x}_t) = \int \max_{a_t \in A_t} \left\{ u(\mathbf{x}_t, a_t) + \epsilon_t(a_t) + \alpha \int \overline{V}_{t+1}(\mathbf{x}_{t+1}) f(\mathbf{x}_{t+1}|\mathbf{x}_t, a_t) d\mathbf{x}_{t+1} \right\} g(\boldsymbol{\epsilon}_t) d\boldsymbol{\epsilon}_t \tag{9}$$

The ex ante value function represents the analyst's estimate of total utility if the decision maker behaves optimally after time period t. This observation enables us to mathematically define the conditional value function as

$$v_t(\mathbf{x}_t, a_t) = u(\mathbf{x}_t, a_t) + \alpha \int \overline{V}_{t+1}(\mathbf{x}_{t+1}) f(\mathbf{x}_{t+1}|\mathbf{x}_t, a_t) d\mathbf{x}_{t+1} \tag{10}$$

Using (10), we can simplify (9) as follows

$$\overline{V}_t(\mathbf{x}_t) = \int \max_{a_t \in A_t} \left\{ v_t(\mathbf{x}_t, a_t) + \epsilon_t(a_t) \right\} g(\boldsymbol{\epsilon}_t) d\boldsymbol{\epsilon}_t \tag{11}$$

When ϵ's are extreme value type I distributed, the ex ante value functions can be further simplified as

$$\overline{V}_t(\mathbf{x}_t) = \gamma + \ln \sum_{a_t \in A_t} \exp\big(v_t(\mathbf{x}_t, a_t)\big) \tag{12}$$

where γ represents the Euler's constant. Thus, using equations (10) and (12), we can construct a backward recursive algorithm to find the conditional and ex ante value functions as summarized in Algorithm 1.

Algorithm 1

Step 1: Initialize Terminal Value Functions

$v_T(\mathbf{x}_T, a_T) = u(\mathbf{x}_T, a_T) \, \forall \, \mathbf{x}_T \in X_T, a_T \in A_T$

$\overline{V}_T(\mathbf{x}_T) = \gamma + \ln \sum_{a_T \in A_T} \exp\big(v_T(\mathbf{x}_T, a_T)\big) \, \forall \, \mathbf{x}_T \in X_T$

Step 2: Backward Recursion

for $t = T-1, T-2, \ldots, 1$ **do**

$v_t(\mathbf{x}_t, a_t) = u(\mathbf{x}_t, a_t) + \alpha \int \overline{V}_{t+1}(\mathbf{x}_{t+1}) f(\mathbf{x}_{t+1}|\mathbf{x}_t, a_t) d\mathbf{x}_{t+1} \, \forall \, \mathbf{x}_t \in X_t, a_t \in A_t$

$\overline{V}_t(\mathbf{x}_t) = \gamma + \ln \sum_{a_t \in A_t} \exp\big(v_t(\mathbf{x}_t, a_t)\big) \, \forall \, \mathbf{x}_t \in X_t$

end for

3.2 Application to Hurricane Evacuation

As mentioned earlier, Atlantic hurricanes can be tracked several days before landfall. During this period, real-time information on hurricanes and forecasts issued by meteorological agencies assist individuals in making evacuation-related choices. These updates are typically relayed every six hours, and each six-hour block constitutes one time period in our study. In each time period, individuals may decide to evacuate or wait, as shown in Figure 1. If individuals choose to evacuate in a particular time period, they do not make any choices in the future periods, but if they choose to wait, they again face two options: evacuate or wait. However, in the final period, individuals can either evacuate or decide to stay (and never evacuate). As is standard practice in evacuation literature, we assume

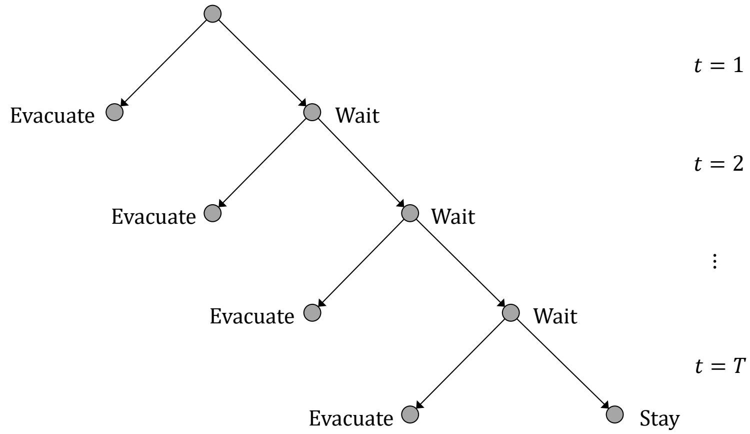

Figure 1: Choices during hurricane evacuation.

that the time to make landfall is known with a greater degree of certainty, and this information dictates the number of time periods in the model.

The observable state vector $\mathbf{x}$ at each time step is assumed to comprise of an intercept 1, socio-demographic variables (such as household size, number of vehicles, and number of years of residence), and time-varying hurricane-related characteristics (such as intensity, distance from household to the center of the hurricane, storm surge, and wind speeds). Forecasts on landfall location, intensity, and wind speeds can also be included in the observable state vector.

Usually, in dynamic discrete choice models, when utilities represent tangible quantities such as cost, it is reasonable to imagine that individuals maximize a discounted sum of per-period or one-step utilities. However, in the context of evacuation, the one-step utilities (or rather disutilities) are not very definitive. One way to interpret them is to suppose that they reflect the anxiety experienced by individuals. In other words, we imagine that individuals, when faced with the option of evacuating, make sequential decisions to minimize their overall anxiety.

The observable components of the one-step utility functions are assumed to be linearly dependent on the covariates. In time periods leading up to the last one, decision makers can choose between evacuating and waiting, both of which are likely to result in varying levels of anxiety. Thus, it is sensible to assume that the impacts of state variables for evacuating are different from that due to waiting. For similar reasons, we suppose that the evacuate and stay options in the last time period result in different utilities. Also, note that choosing to evacuate in any time period terminates the decision-making process. Hence, the one-step utilities associated with the evacuate option do not simply represent the anxiety caused in a six-hour time window but also capture the inconvenience of traveling to a safe location. In other words, evacuation choices yield a one-time terminal utility, and no additional utility is added from subsequent time periods. Similarly, the utility of the stay option in the last time period includes the discomfort caused until normal conditions are restored after a hurricane has passed. The stay option can be viewed as the de facto choice.[3] Specifically, we assume that $u(\mathbf{x}_t, \texttt{Evacuate}) = \boldsymbol{\beta}^\mathsf{T}\mathbf{x}_t$, $u(\mathbf{x}_t, \texttt{Wait}) = \boldsymbol{\psi}^\mathsf{T}\mathbf{x}_t$ and $u(\mathbf{x}_t, \texttt{Stay}) = 0$, where the vectors $\boldsymbol{\beta}$ and $\boldsymbol{\psi}$ are parameters to be estimated from the data. Unlike regular discrete choice models, estimating coefficients for attributes that appear in both evacuate and wait utilities is possible. The stay option's observable utility in the last time step is normalized to zero to avoid identification issues since the state vector does not include any alternative-specific variables. Such a normalization scheme is well suited in other dynamic choice contexts. For example, consider an analogous problem of deciding between renting and buying a house over a person's lifetime. Individuals receive utilities in the intermediate periods when they rent, and choosing to buy a house may be assumed to yield a terminal utility, both of which may depend on a variety of factors such as income, market value, and mortgage rates. If the person does not buy a house in the last period, the associated utility may be set to zero. In theory, for the evacuation problem, one can also choose to normalize the utilities of the other options to zero, but normalizing the stay option resulted in better log-likelihood function values for the data used in this paper. More explanation on the specification of the utilities and associated identification issues can be found in Appendix A.

[3] According to Mileti and Sorensen (1990), evacuation choices are typically motivated by disaster-specific information, which is processed by individuals in multiple stages: hearing, understanding, believing, personalizing, and deciding and responding. Thus, we assume that the stay option will be preferred when such causal triggers are absent or not strong.

4 Maximum Likelihood Estimation

Dynamic choice models are usually estimated using panel data spanning the time period of interest. For the current application, we populate the observable component of the state vector using socio-demographic features and time-dependent hurricane characteristics. Let N represent the set of households and $\mathbf{x}_{nt}$ denote the observable state vector of household n in time period t.

4.1 Perfect Information Models

In perfect information models, we suppose that state transitions take place with certainty. This is limiting in practice because individuals would not know the future trajectory and intensity of the hurricane with certainty at any point in time. Hence, the estimation results from these models are not very informative. However, it is useful to discuss these models for pedagogical reasons since they serve as a good starting point for advanced models and are simpler to estimate. Perfect information models have also been formulated and estimated in earlier papers on hurricane evacuation (Gudishala and Wilmot, 2012; Serulle and Cirillo, 2017).

Since the hurricane characteristics are known for all time periods, the state space of household n, X_{nt}, is a singleton containing the data on observed covariates in period t, which we will denote using $\mathbf{y}_{nt}$ (we will continue to represent generic state variables using $\mathbf{x}$). Thus, the conditional value functions in the backward recursion step (Step 2 of Algorithm 1) can be expressed as [4]

$$v_{nt}(\mathbf{y}_{nt}, \texttt{Evacuate}) = \boldsymbol{\beta}^{\mathsf{T}}\mathbf{y}_{nt} \tag{13}$$

$$v_{nt}(\mathbf{y}_{nt}, \texttt{Wait}) = \boldsymbol{\psi}^{\mathsf{T}}\mathbf{y}_{nt} + \alpha\overline{V}_{n,t+1}(\mathbf{y}_{n,t+1}) \tag{14}$$

Define a vector $\boldsymbol{\lambda} = (\boldsymbol{\beta}, \boldsymbol{\psi}, \alpha)$ which includes all the parameters that may be estimated from the data. Suppose the decisions of the household are represented using an indicator variable $\delta_{nt}(a_t)$ which takes a value 1 if action a_t is selected by household n in period t and is 0 otherwise. Once a household evacuates, $\delta_{nt}(a_t)$ is set to 0 for all future actions. The analyst's estimate of the probability with which household n chooses action a_t for a given vector of parameters $\boldsymbol{\lambda}$ is

$$\mathbb{P}[a_t|\mathbf{y}_{nt}; \boldsymbol{\lambda}] = \frac{\exp\left(v_{nt}(\mathbf{y}_{nt}, a_t; \boldsymbol{\lambda})\right)}{\sum_{a'_t \in A_t} \exp\left(v_{nt}(\mathbf{y}_{nt}, a'_t; \boldsymbol{\lambda})\right)} \tag{15}$$

The role of the parameters $\boldsymbol{\lambda}$ are explicitly highlighted by adding them to the arguments of the choice probability and the conditional value function. The likelihood of observing a household n's sequence of actions or policy is $\prod_{t=1}^{T}\prod_{a_t\in A_t}\mathbb{P}[a_t|\mathbf{y}_{nt};\boldsymbol{\lambda}]^{\delta_{nt}(a_t)}$ and hence the log-likelihood function can be written as

$$\mathcal{LL}(\boldsymbol{\lambda}) = \sum_{n=1}^{N}\sum_{t=1}^{T}\sum_{a_t\in A_t}\delta_{nt}(a_t)\ln\mathbb{P}[a_t|\mathbf{y}_{nt};\boldsymbol{\lambda}] \tag{16}$$

The goal of the estimation procedure is to find $\boldsymbol{\lambda}^*$ which maximizes $\mathcal{LL}(\boldsymbol{\lambda})$, subject to $\alpha \in (0, 1]$.[5] (The remaining parameters are unconstrained.) The constrained nonlinear program can be transformed into an unconstrained optimization problem by replacing the discount factor α with a continuous differentiable function $h(\alpha)$ whose domain is the set of reals $\mathbb{R}$ and whose range is $(0, 1)$. Examples of such functions include $1/(1 + \exp(-\alpha))$ and $\alpha^2/(1 + \alpha^2)$. To check for the optimality of the endpoint, one can simply fix α to 1 and optimize the log-likelihood function by varying the other parameters. The problem can then be solved using a standard quasi-Newton approach such as the BHHH algorithm (Berndt et al., 1974). Note, however, that objective is non-concave, and hence the algorithm can only guarantee a local maxima.

Let $\mathcal{LL}_n(\boldsymbol{\lambda})$ be household n's contribution to the log-likelihood function, i.e.,

$$\mathcal{LL}_n(\boldsymbol{\lambda}) = \sum_{t=1}^{T}\sum_{a_t\in A_t}\delta_{nt}(a_t)\ln\mathbb{P}[a_t|\mathbf{y}_{nt};\boldsymbol{\lambda}] \tag{17}$$

[4] The subscript n in the expressions for the conditional and ex ante value functions $v(.)$ and $V(.)$ are redundant but are retained to emphasize that they are different for different households

[5] Setting α to 0 implies that individuals are myopic, and their choice probabilities can be obtained using a logit equation at each stage. Note that, in this case, the utility equations must be modified to avoid additional identification issues. On the other hand, $\alpha = 1$ indicates that users place equal importance on current and future utilities. This is reasonable in the context of hurricane evacuations since the time period of interest is just a few hours or days and is not in the order of years.

Let $\intercal$ represent the transpose operator. Define a vector of 'scores' $\mathbf{s}_n(\boldsymbol{\lambda})$ for household n as

$$(\mathbf{s}_n(\boldsymbol{\lambda}))^\intercal = \left[\nabla^\intercal_{\boldsymbol{\beta}} \mathcal{LL}_n(\boldsymbol{\lambda}) \quad \nabla^\intercal_{\psi} \mathcal{LL}_n(\boldsymbol{\lambda}) \quad \frac{\partial}{\partial \alpha} \mathcal{LL}_n(\boldsymbol{\lambda})\right] \tag{18}$$

Using (17), the derivatives of household n's contribution to the log-likelihood function can be written as

$$\mathbf{s}_n(\boldsymbol{\lambda}) = \sum_{t=1}^{T} \sum_{a_t \in A_t} \delta_{nt}(a_t) \frac{\nabla_{\boldsymbol{\lambda}} \mathbb{P}[a_t|\mathbf{y}_{nt}; \boldsymbol{\lambda}]}{\mathbb{P}[a_t|\mathbf{y}_{nt}; \boldsymbol{\lambda}]} \tag{19}$$

Details regarding the derivatives of the choice probabilities can be found in Appendix B.1. Denote using $\mathbf{B}_n(\boldsymbol{\lambda})$, the outer product of the household n's scores, i.e.,

$$\mathbf{B}_n(\boldsymbol{\lambda}) = \mathbf{s}_n(\boldsymbol{\lambda}) \otimes \mathbf{s}_n(\boldsymbol{\lambda}) = \mathbf{s}_n(\boldsymbol{\lambda})(\mathbf{s}_n(\boldsymbol{\lambda}))^\intercal \tag{20}$$

Finally, define $\mathbf{s}(\boldsymbol{\lambda})$ and $\mathbf{B}(\boldsymbol{\lambda})$ as the average of the outer products across the sample.

$$\mathbf{s}(\boldsymbol{\lambda}) = \frac{1}{N} \sum_{n=1}^{N} \mathbf{s}_n(\boldsymbol{\lambda}) \tag{21}$$

$$\mathbf{B}(\boldsymbol{\lambda}) = \frac{1}{N} \sum_{n=1}^{N} \mathbf{B}_n(\boldsymbol{\lambda}) \tag{22}$$

The vector $\mathbf{s}(\boldsymbol{\lambda})$ is simply the gradient of the log-likelihood function and the matrix $\mathbf{B}(\boldsymbol{\lambda})$ serves as an approximation of the information matrix at the true parameters (Train, 2009). The $\boldsymbol{\lambda}$s are updated from one iteration k to the next in the following way

$$\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k + \mu_k \left(\mathbf{B}(\boldsymbol{\lambda}_k)\right)^{-1} \mathbf{s}(\boldsymbol{\lambda}_k) \tag{23}$$

where μ_k is the step size that can be computed using a backtracking line search method. One can begin with a step size of 1 and keep halving it until the objective improves. The algorithm is assumed to converge when $(\mathbf{s}(\boldsymbol{\lambda}))^\intercal(\mathbf{B}(\boldsymbol{\lambda}))^{-1}\mathbf{s}(\boldsymbol{\lambda})$ is less than a predetermined threshold.

4.2 Partial Information Models

In partial information models, we suppose that individuals only know the current values of the attributes in each time period. They, however, do have beliefs on the evolution of the covariates. These beliefs stem from either their own past experiences or from forecasted conditions (or sometimes both). In the subsequent subsections, we discuss models for two situations: one in which forecasts are not available (or are not widely publicized) and individuals form their own beliefs, and another in which they fully rely on forecasts from NHC for estimating their future utilities.

4.2.1 Without Forecasts

In the first approach, individuals' beliefs are assumed to follow certain parameterized distributions specified by the analyst. The parameters of these distributions are estimated alongside with the coefficients of the one-step utilities using the maximum likelihood method. Theoretically, one can model individuals' beliefs on the evolution of mandatory orders, hurricane characteristics such as the distance from the center, and even the predictions about the landfall location, wind speed, and storm surge. However, many of these variables are continuous, and to solve the dynamic program, it is necessary to discretize them. Furthermore, the state space grows exponentially with the number of time-varying attributes, which results in the well-known *curse of dimensionality*, making it difficult to solve the dynamic program and also to extract the transition probability parameters. For these reasons, we restrict the observable state to contain only socio-demographic features (which remain constant over time) and discrete hurricane-related characteristics such as intensity. Note that the Markovian transition functions $f(\mathbf{x}_{t+1}|\mathbf{x}_t, a_t)$ do not depend on the actions of households since storm-specific threats are not influenced by evacuation decisions and can hence be written as $f(\mathbf{x}_{t+1}|\mathbf{x}_t)$.

Mathematically, we write the state space of household n at time t, X_{nt}, as $X_{nt} = \{\mathbf{x}^s_{nt}\} \times \{0, 1, 2, 3, 4, 5\}$, where $\times$ represents the Cartesian product and $\mathbf{x}^s_{nt}$ is the vector of socio-demographic attributes and the scalar $x^i_{nt} \in$

$\{0, 1, \ldots, 5\}$ denotes the hurricane intensity.[6] Individuals are assumed to know the current hurricane intensity and have a probabilistic opinion of how it might evolve in the next time period. These beliefs can form from past experiences as well as from external sources that are latent to our specification.[7] The conditional value functions in the backward recursion step (Step 2 of Algorithm 1) can thus be expressed as

$$v_{nt}(\mathbf{y}_{nt}, \texttt{Evacuate}) = \boldsymbol{\beta}^{\mathsf{T}}\mathbf{y}_{nt} \tag{24}$$

$$v_{nt}(\mathbf{y}_{nt}, \texttt{Wait}) = \boldsymbol{\psi}^{\mathsf{T}}\mathbf{y}_{nt} + \alpha \sum_{\substack{\mathbf{x}_{n,t+1} \in \\ X_{n,t+1}}} f(\mathbf{x}_{n,t+1}|\mathbf{y}_{nt};\theta)\, \overline{V}_{n,t+1}(\mathbf{x}_{n,t+1}) \tag{25}$$

where $f(\mathbf{x}_{n,t+1}|\mathbf{x}_{nt};\theta) = \mathbb{P}[x^i_{n,t+1}|x^i_{nt};\theta]$ is the probability with which the intensity in period $t+1$ will be $x^i_{n,t+1}$, given that the intensity in the current period t is x^i_{nt}. We suppose that this transition matrix is parameterized by θ that represents the probability of observing the same intensity in the next time period, and is expected to take the form shown in (26) and the transition diagram 2.

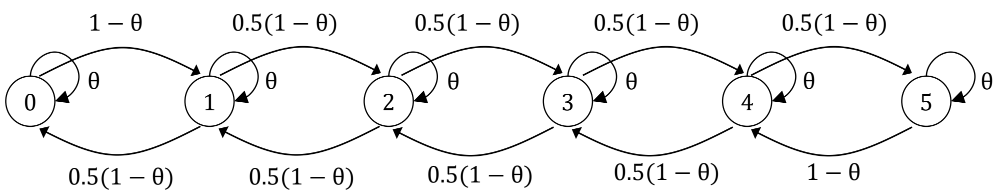

Figure 2: Transition diagram of the assumed belief structure.

If the current intensity is 0 or 5, we assume that the next period's intensity could be 1 or 4, respectively, with probability $1-\theta$. On the other hand, if the current intensity is 1, 2, 3, or 4, the intensity in the next period could take immediate neighboring values with probability $0.5(1-\theta)$.

$$\mathbb{P}[x^i_{n,t+1}|x^i_{nt};\theta] = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} \theta & 1-\theta & 0 & 0 & 0 & 0 \\ 0.5(1-\theta) & \theta & 0.5(1-\theta) & 0 & 0 & 0 \\ 0 & 0.5(1-\theta) & \theta & 0.5(1-\theta) & 0 & 0 \\ 0 & 0 & 0.5(1-\theta) & \theta & 0.5(1-\theta) & 0 \\ 0 & 0 & 0 & 0.5(1-\theta) & \theta & 0.5(1-\theta) \\ 0 & 0 & 0 & 0 & 1-\theta & \theta \end{pmatrix} \end{array} \tag{26}$$

Our choice of the transition matrix was motivated by two major observations. First, hurricane intensity is unlikely to vary drastically in a six-hour period. E.g., the absolute difference in the intensity between consecutive periods for Hurricane Gustav was at most 1. Second, the specification is simple and involves a single parameter which makes the estimation process easier. Note that θ must belong to $[0, 1]$. Alternatively, we can specify different versions of parameterized transition matrices and select one that provides a better fit. However, the resulting maximum likelihood optimization problems can be complicated due to the presence of more variables and constraints.

For the partial information model without forecasts, let $\boldsymbol{\lambda} = (\boldsymbol{\beta}, \boldsymbol{\psi}, \alpha, \theta)$. The time-dependent choice probabilities are exactly the same as in (15). The log-likelihood on the other hand has an extra term associated with state transitions and can be represented as

$$\mathcal{LL}(\boldsymbol{\lambda}) = \sum_{n=1}^{N}\sum_{t=1}^{T}\sum_{a_t \in A_t} \delta_{nt}(a_t)\Big\{ \ln \mathbb{P}[a_t|\mathbf{y}_{nt};\boldsymbol{\lambda}] + \ln f(\mathbf{y}_{nt}|\mathbf{y}_{n,t-1};\theta) \Big\} \tag{27}$$

[6] The index t in $\mathbf{x}^s_{nt}$ is not completely redundant since some household-specific variables such as number of available vehicles can change over time. On the other hand, the index n for x^i_{nt} is redundant but will be retained for uniformity.

[7] In Rust (1987), a single individual made dynamic replacement decisions for multiple engines and hence the function f captured the decision maker's estimates of future mileage. On the other hand, in the case of evacuation, it is possible for different households to have different beliefs of the future. However, extracting parameters from such specifications is very difficult and instead, we assume that the function f does not depend on n and reflects the collective belief of all the households in the sample.

When $t = 1$, the value of intensity in the previous period is set to 0. As was the case in perfect information models, the objective is to find $\boldsymbol{\lambda}^*$ that maximizes $\mathcal{LL}(\boldsymbol{\lambda})$, subject to $\alpha \in (0, 1]$, and $\theta \in [0, 1]$. The constrained variables can similarly be replaced with $h(\alpha)$ and $h(\theta)$ to obtain an unconstrained optimization problem with an open domain. Additionally, we check if the optimal solution occurs at the boundaries by fixing the values of the bounded parameters. One could also just optimize the term associated with the state transitions, $\ln f(\mathbf{y}_{nt}|\mathbf{y}_{n,t-1}; \theta)$, to get an estimate $\hat{\theta}$, fix it, and estimate the remaining parameters. This two-step estimation procedure is computationally efficient but can induce errors since the log-likelihood function is not separable (its first term also depends on θ).

Household n's contribution to the log-likelihood function can analogously be written as

$$\mathcal{LL}_n(\boldsymbol{\lambda}) = \sum_{t=1}^{T} \sum_{a_t \in A_t} \delta_{nt}(a_t) \Big\{ \ln \mathbb{P}[a_t|\mathbf{y}_{nt}; \boldsymbol{\lambda}] + \ln f(\mathbf{y}_{nt}|\mathbf{y}_{n,t-1}; \theta) \Big\} \tag{28}$$

The score vector $\mathbf{s}_n(\boldsymbol{\lambda})$ for household n is defined as

$$(\mathbf{s}_n(\boldsymbol{\lambda}))^{\mathsf{T}} = \left[\nabla^{\mathsf{T}}_{\boldsymbol{\beta}} \mathcal{LL}_n(\boldsymbol{\lambda}) \;\; \nabla^{\mathsf{T}}_{\psi} \mathcal{LL}_n(\boldsymbol{\lambda}) \;\; \frac{\partial}{\partial \alpha} \mathcal{LL}_n(\boldsymbol{\lambda}) \;\; \frac{\partial}{\partial \theta} \mathcal{LL}_n(\boldsymbol{\lambda}) \right] \tag{29}$$

and can be computed using

$$\mathbf{s}_n(\boldsymbol{\lambda}) = \sum_{t=1}^{T} \sum_{a_t \in A_t} \delta_{nt}(a_t) \left\{ \frac{\nabla_{\boldsymbol{\lambda}} \mathbb{P}[a_t|\mathbf{y}_{nt}; \boldsymbol{\lambda}]}{\mathbb{P}[a_t|\mathbf{y}_{nt}; \boldsymbol{\lambda}]} + \frac{\nabla_{\boldsymbol{\lambda}} f\big(\mathbf{x}_{nt}|\mathbf{x}_{n,t-1}; \theta\big)}{f\big(\mathbf{x}_{nt}|\mathbf{x}_{n,t-1}; \theta\big)} \right\} \tag{30}$$

The outer product and BHHH algorithm outlined in (20) - (23) can then be used to find the optimal $\boldsymbol{\lambda}^*$. Calculations of the derivatives of the objective with respect to the decision variables are detailed in Appendix B.2.1.

4.2.2 With Forecasts

In a second approach, we suppose that individuals' beliefs on the evolution of the hurricane are derived from forecasts made by the NHC. The NHC provides forecasts of the intensity, distance to landfall, and wind speeds at select geographical locations every four to six hours. These forecasts can be used to construct the transition probability matrices of the system states. However, a key difference from the previous models is that the transition functions are non-stationary, i.e., they vary over time. Hence, at each time step t, the analyst estimates the ex ante value function for all future time steps.

Mathematically, one can write the state space of household n at time t, X_{nt}, as $X_{nt} = \{\mathbf{x}^s_{nt}\} \times \{0, 1, 2, 3, 4, 5\} \times \mathbb{R}^+$, where $\times$ represents the Cartesian product and $\mathbf{x}^s_{nt}$ is the vector of socio-demographic attributes, the scalar $x^i_{nt} \in \{0, 1, \ldots, 5\}$ denotes the hurricane intensity, and $x^d_{nt} \in \mathbb{R}^+$ represents the distance between household n and the center of the storm. A snapshot of predictions made by NHC is shown in Figure 3.

- - - MAXIMUM WIND SPEED (INTENSITY) PROBABILITIES - - -

VALID TIME	06Z SAT	18Z SAT	06Z SUN	18Z SUN	18Z MON	18Z TUE	18Z WED
FORECAST HOUR	12	24	36	48	72	96	120
DISSIPATED	X	X	X	X	3	10	33
TROP DEPRESSION	X	1	X	X	5	11	20
TROPICAL STORM	9	10	5	2	16	27	24
HURRICANE	91	89	95	98	76	52	23
HUR CAT 1	79	34	18	12	20	20	8
HUR CAT 2	10	38	30	21	21	13	5
HUR CAT 3	2	16	38	37	22	12	6
HUR CAT 4	X	2	8	23	11	6	4
HUR CAT 5	X	1	1	4	2	1	1
FCST MAX WIND	75KT	90KT	100KT	110KT	105KT	85KT	55KT

FORECAST POSITIONS AND MAX WINDS

INITIAL	29/1500Z	18.6N	78.8W	55 KT
12HR VT	30/0000Z	19.3N	80.0W	70 KT
24HR VT	30/1200Z	20.8N	81.8W	80 KT
36HR VT	31/0000Z	22.6N	83.8W	95 KT...OVER WESTERN CUBA
48HR VT	31/1200Z	24.3N	85.7W	105 KT
72HR VT	01/1200Z	27.5N	89.0W	105 KT
96HR VT	02/1200Z	29.5N	91.0W	100 KT...INLAND
120HR VT	03/1200Z	31.0N	93.0W	60 KT...INLAND

Figure 3: Forecasts on Hurricane Gustav from NHC archives.
(Source: www.nhc.noaa.gov/archive/2008/)

Note from Figure 3 that the intensity forecasts are in the form of a probability distribution, whereas the position forecasts are deterministic. Let $f_{tt'}(\mathbf{x}_{nt'})$ denote the probability mass function at state $\mathbf{x}_{nt'}$, where $t' > t$, as predicted

Algorithm 2

Step 1: Initialize Terminal Value Functions
$v_{nT}(\mathbf{y}_{nT}, a_T) = u_n(\mathbf{y}_{nT}, a_T)\,\forall\, a_T \in A_T = \{\texttt{Evacuate}, \texttt{Stay}\}$
$\overline{V}_{ntT}(\mathbf{x}_{nT}) = \gamma + \ln \sum_{a_T \in A_T} \exp\left(v_{nT}(\mathbf{x}_{nT}, a_T)\right) \forall\, \mathbf{x}_{nT} \in X_{nT}, t \in \{1, 2, \ldots, T-1\}$

Step 2: Backward Recursion
for $t = 1, 2, \ldots, T-1$ **do**
  for $t' = T-1, T-2, \ldots, t+1$ **do**
    $$\overline{V}_{ntt'}(\mathbf{x}_{nt'}) = \gamma + \ln \Big\{ \exp(\boldsymbol{\beta}^\mathsf{T}\mathbf{x}_{nt'}) + \exp\Big(\boldsymbol{\psi}^\mathsf{T}\mathbf{x}_{nt'} + \alpha \sum_{\substack{\mathbf{x}_{n,t'+1} \in \\ X_{n,t'+1}}} f_{tt'}(\mathbf{x}_{n,t'+1}) \overline{V}_{nt,t'+1}(\mathbf{x}_{n,t'+1})\Big)\Big\} \forall\, \mathbf{x}_{nt'} \in X_{nt'}$$
  end for
  $v_{nt}(\mathbf{y}_{nt}, \texttt{Evacuate}) = \boldsymbol{\beta}^\mathsf{T}\mathbf{y}_{nt}$
  $v_{nt}(\mathbf{y}_{nt}, \texttt{Wait}) = \boldsymbol{\psi}^\mathsf{T}\mathbf{y}_{nt} + \alpha \sum_{\substack{\mathbf{x}_{n,t+1} \in \\ X_{n,t+1}}} f_{t,t+1}(\mathbf{x}_{n,t+1})\, \overline{V}_{nt,t+1}(\mathbf{x}_{n,t+1})$
end for

by the NHC at time step t.[8]

As before, assume that $v_{nt}(\mathbf{x}_{nt}, a_t)$ is the conditional value function that provides an estimate of the utility received from choosing a_t in period t and behaving optimally thereafter. However, the vs are calculated using the following modified Bellman equations:

$$v_{nt}(\mathbf{y}_{nt}, \texttt{Evacuate}) = \boldsymbol{\beta}^\mathsf{T}\mathbf{y}_{nt} \tag{31}$$

$$v_{nt}(\mathbf{y}_{nt}, \texttt{Wait}) = \boldsymbol{\psi}^\mathsf{T}\mathbf{y}_{nt} + \alpha \sum_{\substack{\mathbf{x}_{n,t+1} \in \\ X_{n,t+1}}} f_{t,t+1}(\mathbf{x}_{n,t+1})\, \overline{V}_{nt,t+1}(\mathbf{x}_{n,t+1}) \tag{32}$$

The ex ante value functions also cannot be directly obtained from (12) because the transition probabilities are time-dependent. They are instead calculated for each time step t using backward recursion as shown in (33).

$$\overline{V}_{ntt'}(\mathbf{x}_{nt'}) = \gamma + \ln \left\{ \exp(\boldsymbol{\beta}^\mathsf{T}\mathbf{x}_{nt'}) + \exp\Big(\boldsymbol{\psi}^\mathsf{T}\mathbf{x}_{nt'} + \alpha \sum_{\substack{\mathbf{x}_{n,t'+1} \in \\ X_{n,t'+1}}} f_{tt'}(\mathbf{x}_{n,t'+1}) \overline{V}_{nt,t'+1}(\mathbf{x}_{n,t'+1})\Big) \right\} \tag{33}$$

where t' is a time step greater than t. The steps for calculating the conditional and ex ante value functions for a household n are summarized in Algorithm 2.

The optimal parameters and the discount factor can be estimated using equations (15)-(23) as discussed in Section 4.1. However, since the ex ante value functions are calculated differently, the formulae for the derivatives of the log-likelihood objective are slightly different and are described in Appendix B.2.2.

The current paper differs from Serulle and Cirillo (2017) in three important ways. First, we observe that the evacuation problem is a finite-horizon model with terminal choices. Hence, the one-step utilities of waiting are distinguished from the utility of the stay option. Serulle and Cirillo (2017), on the other hand, set the utilities of the wait option to zero as well, which imposes more restrictions on the model structure. Second, in the partial information models, we assume that the transition probabilities of the system states are computed only using data on Hurricane Gustav and not from previous Atlantic hurricanes (since most of them did not affect the households in this study). We model two variants of the problem, one with a parameterized stationary belief system and another that uses non-stationary NHC data.[9] Finally, we do not approximate the value functions but solve the complete

[8] The transition function f can be indexed by n if we have predictions on the probability of wind speeds for different households or parishes, but such information was not available.

[9] Using data from past hurricanes may improve transferability, but survey data from households affected by different hurricanes would be required to validate this guess.

dynamic program using utilities from all future time steps.[10] Additionally, we provide a thorough exposition on the interpretation of signs of model parameters in Section 5.3 using time-dependent sensitivities.

5 Results

5.1 Data Description

In this section, we present the estimation results of the dynamic discrete choice model using household survey data from Hurricane Gustav that was collected by the Public Policy Research Lab at Louisiana State University (Wilmot et al., 2013). The survey questionnaire was designed to gather information on socio-demographic attributes of households, residential area-specific characteristics, evacuation decisions and reasons, vehicle ownership, and destination choices. Some of the missing values in the data were imputed using the software R. The time horizon for the dynamic choice model was set from 12:00 AM, 28 August 2008 to 12:00 AM, 1 September 2008, and contained 16 six-hour time periods. Voluntary evacuations orders were issued in several parishes on August 30, and a state-wide mandatory evacuation order was declared at 8 AM on August 31. Figure 4 shows the surveyed households and the landfall location. The household data set contained 277 households, of which 193 households evacuated. The cumulative distribution of evacuations over time is shown in Figure 5.

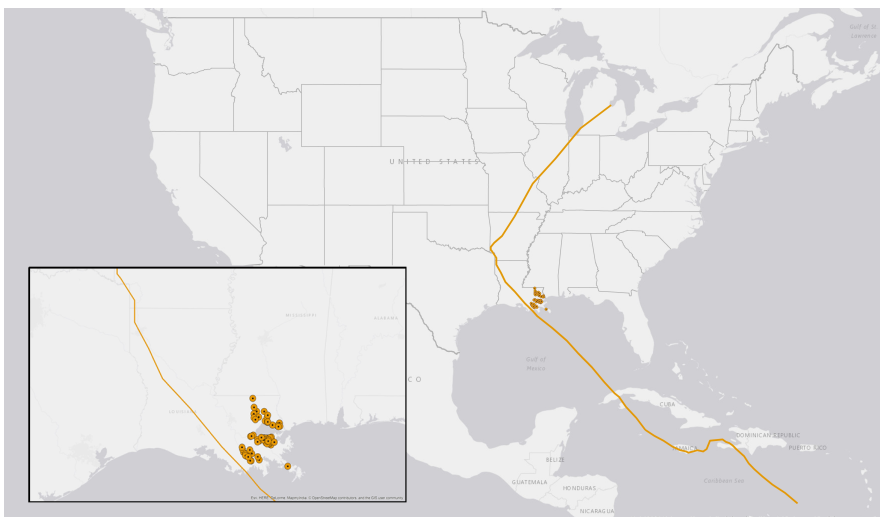

Figure 4: Track of Hurricane Gustav and the locations of surveyed households.

Hurricane Gustav made landfall in Louisiana at 9:30 AM CDT on 1 September 2008. Details regarding its track and intensity can be found in Beven and Kimberlain (2009). Gustav was a Category 4 hurricane when it made landfall in Cuba on August 30. It later weakened to a Category 2 storm before hitting the coast of Louisiana. The survey data is supplemented using hurricane data on current location and intensity and forecasted landfall location. This data was obtained from the NHC website `http://www.nhc.noaa.gov/gis/` and was generated using their Sea, Lake, and Overland Surge from Hurricanes (SLOSH) model. Future forecasts for some of these variables were, however, unavailable. In some cases, the timing of the predictions did not align with the time periods in the case study. Probability mass functions of the intensity and location of the storm's center were interpolated in such cases. A brief description of the attributes used in the three models is given in Table 1.

[10] Approximations are unavoidable in scenarios where continuous covariates have probabilistic forecasts. In such cases, one could discretize the state space or calculate the expectation of the ex ante value functions using simulation. However, one must be wary of estimation errors due to approximating future utilities and choice probabilities. See Appendix C for an example that demonstrates this issue.

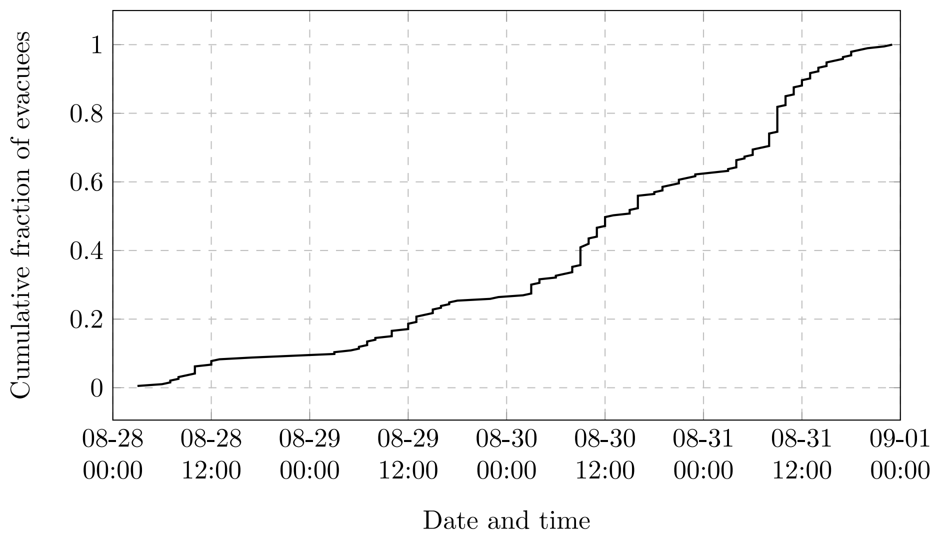

Figure 5: Cumulative demand of evacuees from sampled households over time.

Table 1: Description of covariates used to estimate evacuation choices.

Variable	*Description*
Num Veh	Number of vehicles owned by the household
HH Size	Size of the household (including adults and children)
Income	Total household income in USD (,000s). Income was reported in intervals and the mean values were used in the estimation procedure
Num Seventeen	Number of individuals in the household who are 17 or younger
Have Pets	Binary variable which is set to 1 if the household had pets and 0 otherwise
Schooling Level	Binary variable which is set to 1 if the survey respondent has a college degree and is 0 otherwise
Years Residency	Numbers of year resided at the present household
Distance Center	Distance between the household and the center of the hurricane (in km)
Distance Forecast	Distance between the household and the predicted landfall location (in km)
Intensity	Saffir–Simpson hurricane wind scale (1-5). 0 indicates a tropical depression or a tropical storm
Prob Wind	Probability with which average wind speeds are greater than 74 mph over the next five days
Prob Surge	Probability that the storm surge is greater than 5 ft for the next 80 hours
Mandatory Order	Binary variable which is set to 1 if a state-wide mandatory order was issued

5.2 Estimation Results

The maximum likelihood estimation and the BHHH algorithm were implemented in C++. In the perfect information models, we noticed that when the discount factor was optimized with the remaining parameters, the approximate Hessian turned out to be ill-conditioned and non-invertible because the unconstrained discount factor was being driven to ∞ (which implies that the transformed discount factor $h(\alpha)$ tends to 1). We also verified using multiple fixed discount factor values that the log-likelihood was maximized when it was set to 1.[11]

Table 2 shows the estimation results and the z-values for two specifications of the perfect information model. In Specification I, all attributes were included, and the least significant predictor was iteratively discarded to obtain Specification II, in which only significant attributes were retained. The threshold for convergence of the BHHH algorithm was set to 1E−5. As mentioned earlier, the perfect information model results are not very informative since it relies on a strong assumption that households know everything about the storm in advance.

[11]Similar observations were made in Rust (1987).

Table 2: Estimation results for the perfect information case.

Attribute	Specification I				Specification II			
	β	z-value	ψ	z-value	β	z-value	ψ	z-value
Intercept	7.0620	1.976	−5.5280	−2.455	4.2120	2.788	−5.9170	−5.518
Num Veh	−0.1615	−0.800	0.0304	1.023	-	-	0.0461	2.408
HH Size	−0.0912	−0.352	−0.0244	−1.007	-	-	−0.0218	−1.902
Income	−0.0001	−0.009	0.0005	0.856	-	-	-	-
Num Seventeen	−0.1578	−0.475	−0.0129	−0.385	-	-	-	-
Have Pets	−0.6481	−1.572	−0.0600	−1.262	-	-	-	-
Schooling Level	0.6522	1.626	0.0271	0.589	-	-	-	-
Years Residency	−0.0411	−2.757	−0.0006	−0.390	−0.0325	−3.158	-	-
Distance Center	−0.0154	−2.397	0.0022	1.946	−0.0138	−3.900	0.0025	5.401
Distance Forecast	−0.0221	−3.371	0.0001	0.102	−0.0195	−5.201	-	-
Intensity	0.1080	0.430	−0.0752	−0.293	-	-	-	-
Prob Wind	−0.1938	−1.056	0.0018	0.019	-	-	-	-
Prob Surge	−0.0815	−0.653	0.0206	0.539	-	-	-	-
Mandatory Order	1.9660	2.430	0.4517	0.450	2.0970	5.382	0.5808	1.886
Discount Factor	1.00				1.00			
Log-likelihood	−612.54				−620.50			

The estimation results for the partial information case without forecasts are shown in Table 3. Here again, we started with Specification I that contained all attributes and iteratively excluded the least statistically significant attribute to obtain Specification II and used the same convergence threshold of 1E−5. The optimal discount factor was found to be 1. The parameter θ that is used to describe beliefs on the transition probability matrix was found to be 0.6621 and is statistically significant. Observe that the log-likelihood objective is lower compared to the perfect information case since it also includes the likelihood of observing the intensity transitions as explained in (27).

Table 3: Estimation results for the partial information without forecasts case.

Attribute	Specification I				Specification II			
	β	z-value	ψ	z-value	β	z-value	ψ	z-value
Intercept	−3.6460	−2.542	−0.7891	−4.649	−4.5340	−3.441	−0.9064	−5.185
Num Veh	−0.2778	−1.100	0.0202	0.701	−0.4063	−2.338	-	-
HH Size	0.0533	0.194	−0.0214	−0.864	-	-	−0.0217	−2.112
Income	0.0023	0.339	0.0005	0.780	-	-	0.0006	1.690
Num Seventeen	−0.2268	−0.569	−0.0104	−0.287	-	-	-	-
Have Pets	−0.5424	−1.134	−0.0386	−0.827	-	-	-	-
Schooling Level	0.1108	0.230	0.0275	0.611	-	-	-	-
Years Residency	−0.0305	−1.729	−0.0014	−0.845	−0.0194	−1.757	-	-
Intensity	1.4680	2.506	0.1035	1.834	1.8870	3.048	0.1431	2.498
θ (z-value)	0.6621 (18.390)				0.6621 (18.260)			
Discount Factor	1.00				1.00			
Log-likelihood	−3440.55				−3442.51			

The standard errors for constrained variables like θ and α can be obtained by the delta method or by first solving the transformed problem with $h(\theta)$ or $h(\alpha)$ and then by solving one iteration of the maximum likelihood estimation after initializing the problem with the values shown in the table. We finally estimated the parameters of the partial information model with NHC forecasts. The results are shown in Table 4. The discount factor, in this case, was found to be less than one and significant.

Since the estimated discount factors are not close to zero, the dynamic model can be assumed to provide a better fit compared to a binary logit model. We also estimated the parameters of another static mixed logit model in which individuals either evacuate or stay, but the actions taken in different time periods were treated as different choice occasions for each household. In this model, panel effects were captured using a random parameter for the intercept of the evacuate option. That is, household n's utility of evacuating in choice occasion t (assuming that the household has not evacuated in time period $t-1$ or before) is given by $u_{nt}(\texttt{Evacuate}) = \beta_0^n + \boldsymbol{\beta}^\mathsf{T}\mathbf{y}_{nt}$, where β_0^n is the

Table 4: Estimation results for the partial information with forecasts case.

Attribute	Specification I				Specification II			
	β	z-value	ψ	z-value	β	z-value	ψ	z-value
Intercept	12.7900	2.768	−4.4650	−1.582	7.0268	4.575	−7.4088	−8.781
Num Veh	−0.1469	−0.224	0.0538	0.183	-	-	0.0858	2.707
HH Size	−0.4641	−0.726	−0.2634	−0.972	−0.4679	−2.080	−0.1315	−2.878
Income	0.0006	0.029	0.0011	0.127	-	-	-	-
Num Seventeen	−0.4324	−0.443	−0.1695	−0.415	-	-	-	-
Have Pets	1.0960	0.864	0.5132	0.924	-	-	-	-
Schooling Level	−0.5789	−0.527	−0.2544	−0.545	-	-	-	-
Years Residency	−0.0871	−2.143	−0.0334	−1.646	−0.0587	−2.799	−0.0102	−2.432
Distance Center	−0.0451	−4.225	−0.0134	−1.704	−0.0258	−7.966	-	-
Intensity	3.0830	4.369	1.6580	2.921	1.5880	5.099	0.6243	5.036
Discount Factor	0.6093 (6.206)				0.8649 (79.412)			
Log-likelihood	−627.30				−632.45			

random intercept whose realization is different across households and $\mathbf{y}_{nt}$ denotes the attributes of the hurricane and the household n in time period t. We assume that the $\mathbf{y}_{nt}$ vector excludes the element 1, which was used earlier to model intercepts but includes a covariate that represents the number of elapsed time periods (NETP). To incorporate some knowledge of the forecasted conditions, we also introduced an additional covariate that represents the closest predicted distance between the household and the center of the storm for each choice occasion. These variables were included to make the information available to the dynamic model comparable to that used by the static one.[12] We did not assume random parameters for the other coefficients since the parameters in the dynamic model are assumed constant across the population. The utility of the stay option $u_{nt}(\texttt{Stay})$ was normalized to zero. Note that the dynamic model uses future forecasts in the transition probabilities.[13] The estimation of the static model was carried out using the Apollo package in R (Hess and Palma, 2019) and the simulated log-likelihood was constructed using 10,000 Halton draws.

Table 5: Estimation results of a static model with panel effects.

Attribute	Specification I		Specification II	
	β	z-value	β	z-value
Intercept	89.8602	4.595	91.6556	4.813
Standard Deviation	1.5674	2.007	1.4298	1.927
Num Veh	−0.3739	−1.929	−0.3870	−2.053
HH Size	0.2200	1.316	0.2031	1.791
Income	−0.0030	−0.736	-	-
Num Seventeen	−0.0073	−0.035	-	-
Have Pets	−0.1204	−0.443	-	-
Schooling Level	0.0106	0.036	-	-
Years Residency	−0.0088	−0.834	-	-
Distance Center	−0.0564	−4.985	−0.0571	−5.173
Intensity	0.5578	0.985	-	-
Closest Predicted Distance	0.0390	3.859	0.0380	3.901
Num of Elapsed Time Periods (NETP)	−4.8978	−4.265	−5.1219	−4.625
NETP × Distance Center	0.0020	4.863	0.0021	5.264
NETP × Intensity	−0.1201	−2.376	−0.0748	−3.681
NETP × Closest Predicted Distance	−0.0023	−3.140	−0.0022	−3.090
Log-likelihood	−645.30		−646.50	

[12]Although this specification has time variables, we refer to this as a static framework since the utilities are not summed over time, and the estimation procedure does not require solving dynamic programs.

[13]Forecast variables are not introduced as state variables or covariates in the dynamic model since the beliefs on these forecasts will then have to be codified in the transition probabilities.

The NETP variable, however, was not significant, and the resulting model had a lower log-likelihood and performed poorly on the leave-one-out-cross validation tests compared to the dynamic models. A thorough statistical comparison of the static and dynamic frameworks is difficult because one of them is not a restricted version of the other. We then tested a few specifications in which NETP was interacted with other explanatory variables. One such specification in which the NETP variable was statistically significant and had a better log-likelihood is shown in Table 5. We also tested interactions between NETP and static variables such as the household size and number of vehicles, but they were not statistically significant.

The signs of the estimated parameters, especially those associated with the dynamic variables, conform to intuition and can be used to interpret the effects of different covariates on the evacuation probabilities. Since hurricane-related attributes are dynamic, the choice probabilities also vary with time. However, the marginal effects that reflect the changes in these probabilities due to changes in the covariates have to be derived, keeping the interaction terms in mind. We will discuss these calculations in greater detail in the following section and also highlight the advantages and disadvantages of the static model.

5.3 Time-varying Sensitivities

Typically, in static models, the signs of parameters can be used to understand the effect of covariates on the probability of different choices. However, in dynamic choice models, this holds only for certain types of model specifications. For the utility functions defined in the current paper, one cannot simply use the signs to infer the effect of the covariates. To see why, consider the derivatives of the conditional value functions with respect to household size. Let y^h_{nt} denote the size of household n. The evacuation probability derivative with respect to the household size is given by

$$\frac{\partial \mathbb{P}[\texttt{Evacuate}|\mathbf{y}_{nt}]}{\partial y^h_{nt}} = \frac{-\Big(\exp\big(v_{nt}(\mathbf{y}_{nt}, \texttt{Wait}) - v_{nt}(\mathbf{y}_{nt}, \texttt{Evacuate})\big)\Big)}{\Big(1+\exp\big(v_{nt}(\mathbf{y}_{nt}, \texttt{Wait}) - v_{nt}(\mathbf{y}_{nt}, \texttt{Evacuate})\big)\Big)^2} \times \Big(\frac{\partial}{\partial y^h_{nt}} v_{nt}(\mathbf{y}_{nt}, \texttt{Wait}) - \frac{\partial}{\partial y^h_{nt}} v_{nt}(\mathbf{y}_{nt}, \texttt{Evacuate})\Big) \tag{34}$$

Since the conditional value functions vary over time, the sign and magnitude of the sensitivities of the choice probabilities can also change over time. Note that the derivative estimate in a particular time period measures the change in the evacuation probability, assuming that the household did not evacuate in prior periods and one of the covariates is perturbed while keeping everything else the same (including other time-varying hurricane beliefs or forecasts). The calculations use the information on forecasts at the current time period but do not use any realized hurricane characteristics from future periods.

5.3.1 Without Forecasts

In the model without forecasts, the partial derivatives of the conditional value functions can be expressed as

$$\frac{\partial}{\partial y^h_{nt}} v_{nt}(\mathbf{y}_{nt}, \texttt{Evacuate}) = \beta_h \tag{35}$$

$$\frac{\partial}{\partial y^h_{nt}} v_{nt}(\mathbf{y}_{nt}, \texttt{Wait}) = \psi_h + \alpha \sum_{\substack{\mathbf{x}_{n,t+1}\in \\ X_{n,t+1}}} f(\mathbf{x}_{n,t+1}|y_{nt};\theta)\, \frac{\partial}{\partial y^h_{nt}} \overline{V}_{n,t+1}(\mathbf{x}_{n,t+1}) \tag{36}$$

where β_h and ψ_h represent the parameters associated with the household size attribute. Thus, the sign of $\frac{\partial \mathbb{P}[\texttt{Evacuate}|\mathbf{y}_{nt}]}{\partial y^h_{nt}}$ depends on

$$\underbrace{\beta_h - \psi_h}_{\text{Term I}} - \overbrace{\alpha \sum_{\substack{\mathbf{x}_{n,t+1}\in \\ X_{n,t+1}}} f(\mathbf{x}_{n,t+1}|y_{nt};\theta)\, \frac{\partial}{\partial y^h_{nt}} \overline{V}_{n,t+1}(\mathbf{x}_{n,t+1})}^{\text{Term II}} \tag{37}$$

While the first term in the above expression is constant across time steps (except for the last period due to the stay option), the second term is time dependent. The derivatives in the second term can be computed using backward recursion just as done in Section 4.2.1 and Appendix B.2.1.

Figure 6 shows the evacuation probability's sensitivity to the significant attributes, i.e., household size, number of vehicles, number of years of residency, and household income. The vertical axis on the first two plots is the difference in the evacuation probabilities due to a unit increase in the household size and number of vehicles (since they are discrete quantities). The x-axis indicates the time steps at which the derivatives are estimated. Note that the time-variation in the derivatives for each household is not solely from a change in the perturbed covariate. We fix every other state variable to what was observed in the data at each time period, and hence the time-varying hurricane-specific beliefs and forecasts influence these marginals.

Each plot demonstrates the sensitivities of the choice probabilities for three households: one that does not evacuate, one that evacuates in time-period 6, and one that evacuates in time-period 11. For example, an extra member in the household increases the likelihood of evacuation, especially during the intermediate time periods. Likewise, as income increases, households are less likely to evacuate, especially during early time periods, but the variable appears to have no effect after a certain threshold. One possible reason could be that high-income individuals might have sturdier homes that can withstand a storm.

As expected, the household that evacuated early appears to be more sensitive to changes in the influencing covariate. The variation in the sign is similar across households for all time steps, but the magnitudes differ. The sign of the sensitivities does not change with time in this case but notice that it need not match the sign of the estimated parameter. For instance, the evacuation utility parameter for years of residency was found to be -0.0194 (see Table 3), but from Figure 6, a small increase in the number of years of residency has a positive effect on evacuation.

5.3.2 With Forecasts

The procedure for calculating sensitivities for the model with forecasts also uses (34) but the partial derivatives of the conditional value functions (with respect to, say, the household size) are computed from the following equations.

$$\frac{\partial}{\partial y_{nt}^h} v_{nt}(\mathbf{y}_{nt}, \texttt{Evacuate}) = \beta_h \tag{38}$$

$$\frac{\partial}{\partial y_{nt}^h} v_{nt}(\mathbf{y}_{nt}, \texttt{Wait}) = \psi_h + \alpha \sum_{\substack{\mathbf{x}_{n,t+1}\in \\ X_{n,t+1}}} f_{t,t+1}(\mathbf{x}_{n,t+1}) \frac{\partial}{\partial y_{nt}^h} \overline{V}_{nt,t+1}(\mathbf{x}_{n,t+1}) \tag{39}$$

Thus, the sign of $\frac{\partial \mathbb{P}[\texttt{Evacuate}|\mathbf{y}_{nt}]}{\partial y_{nt}^h}$ depends on

$$\underbrace{\beta_h - \psi_h}_{\text{Term I}} - \overbrace{\alpha \sum_{\substack{\mathbf{x}_{n,t+1}\in \\ X_{n,t+1}}} f_{t,t+1}(\mathbf{x}_{n,t+1}) \frac{\partial}{\partial y_{nt}^h} \overline{V}_{nt,t+1}(\mathbf{x}_{n,t+1})}^{\text{Term II}} \tag{40}$$

where the derivative calculations are again similar to the steps used in estimating the parameters of the model with forecasts.

The calculations of the two terms associated with the household size for a sample household are shown in Table 6. In this case, the sign of the sensitivity expression was found to change with time. The sensitivity of choice probabilities to other attributes is shown in Figure 7. Notice that the probability of evacuation decreases with an increase in household size for later time periods, but when the hurricane is far, the probability of evacuation increases with an increase in household size. A possible reason for this behavior could be that the preparation time and risk in evacuating are likely to be greater for larger households, and when the storm is near, they may be averse to evacuating. Similar trends were observed when the effect of the number of years of residency was analyzed. The number of vehicles had a negative effect on the probability of evacuating. An extra vehicle can add to logistical difficulties since it needs an additional driver, and household members can no longer travel together. The sudden change in the sensitivity at time periods 14 and 15 in some cases is probably because the hurricane was forecasted to weaken and its intensity reduced to a Category 2 storm at time step 15.

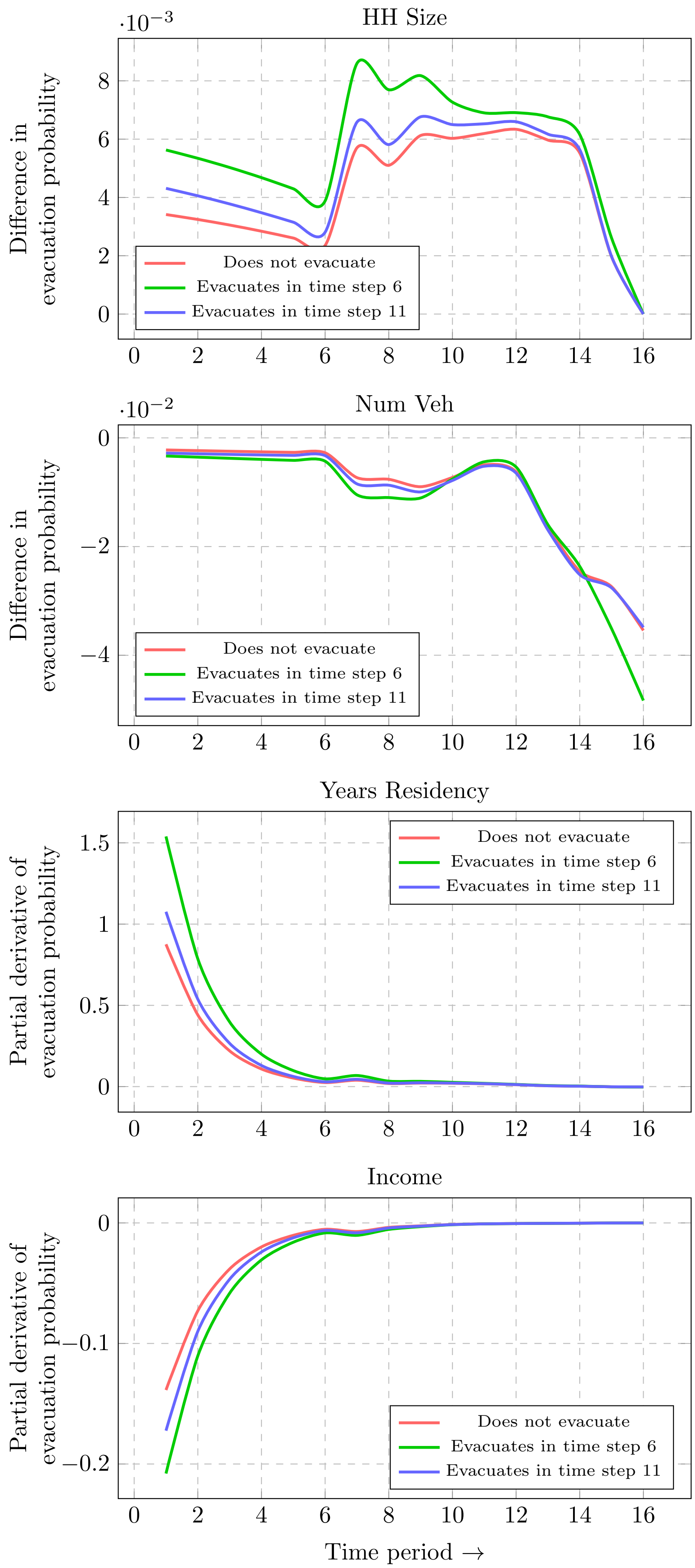

Figure 6: Sensitivity of evacuation probabilities for the model without forecasts.

Table 6: Sensitivity of choice probabilities with respect to household size.

Time Period	β	ψ	Term II	$\beta - \psi -$Term II
16	-0.4679	-	-	-0.4679
15	-0.4679	-0.1315	-0.1267	-0.2097
14	-0.4679	-0.1315	-0.2220	-0.1144
13	-0.4679	-0.1315	-0.3464	0.0100
12	-0.4679	-0.1315	-0.4029	0.0665
11	-0.4679	-0.1315	-0.4371	0.1006
10	-0.4679	-0.1315	-0.4511	0.1147
9	-0.4679	-0.1315	-0.5060	0.1696
8	-0.4679	-0.1315	-0.5604	0.2240
7	-0.4679	-0.1315	-0.5891	0.2527
6	-0.4679	-0.1315	-0.6154	0.2790
5	-0.4679	-0.1315	-0.6406	0.3042
4	-0.4679	-0.1315	-0.6573	0.3209
3	-0.4679	-0.1315	-0.6827	0.3463
2	-0.4679	-0.1315	-0.6848	0.3484
1	-0.4679	-0.1315	-0.6012	0.2647

The sensitivity expression for the distance from the center variable, on the other hand, does not have the second term since the future forecasts do not depend on minor variations in the current distance. Hence, the sign will remain consistent across time periods, but the magnitude of the partial derivatives of the choice probabilities are scaled by the conditional value functions (see equation (34)). Since the intensity values do not change across the population, its coefficients cannot be used to infer its effect on the choice probabilities. The intensity parameters can be viewed as a modification to the intercept term and can be imagined to capture the average time-varying effect of the unobserved factors on the choice preferences.

5.3.3 Static Model

The static framework can also be used to analyze the time-varying sensitivities of the explanatory variables. From Specification II in Table 5, ignoring the time and household indices, the utility expression of the evacuate option for each choice occasion can be written as

$$\begin{aligned} u(\texttt{Evacuate}) = \tilde{\beta}_0 + \beta_1 \text{Num Veh} + \beta_2 \text{HH Size} + \\ \beta_3 \text{Distance Center} + \beta_4 \text{Closest Pred Dist} + \beta_5 \text{NETP} + \\ \beta_6 (\text{NETP})(\text{Distance Center}) + \beta_7 (\text{NETP})(\text{Intensity}) + \\ \beta_8 (\text{NETP})(\text{Closest Pred Dist}) \end{aligned} \quad (41)$$

The expression for the marginals of a variable that does not have an interaction term such as household size y_{nt}^h has a scaling term which is the product of the evacuate and stay choice probabilities multiplied by the coefficient of the variable as shown in (42).

$$\frac{\partial \mathbb{P}[\texttt{Evacuate}]}{\partial y_{nt}^h} = \frac{\exp(-\boldsymbol{\beta}^\mathsf{T} \mathbf{y}_{nt})}{(1 + \exp(-\boldsymbol{\beta}^\mathsf{T} \mathbf{y}_{nt}))^2} \beta_1 \quad (42)$$

Marginals for variables that have an interaction term with the number of time periods elapsed may have two components. For example, suppose y_{nt}^d represents the distance to the storm center for household n at time t. The marginal includes the scaling term, the coefficient of the distance to the center variable β_3, and β_6NETP, which comes from the interaction term.

$$\frac{\partial \mathbb{P}[\texttt{Evacuate}]}{\partial y_{nt}^d} = \frac{\exp(-\boldsymbol{\beta}^\mathsf{T} \mathbf{y}_{nt})}{(1 + \exp(-\boldsymbol{\beta}^\mathsf{T} \mathbf{y}_{nt}))^2} (\beta_3 + \beta_6 \text{NETP}) \quad (43)$$

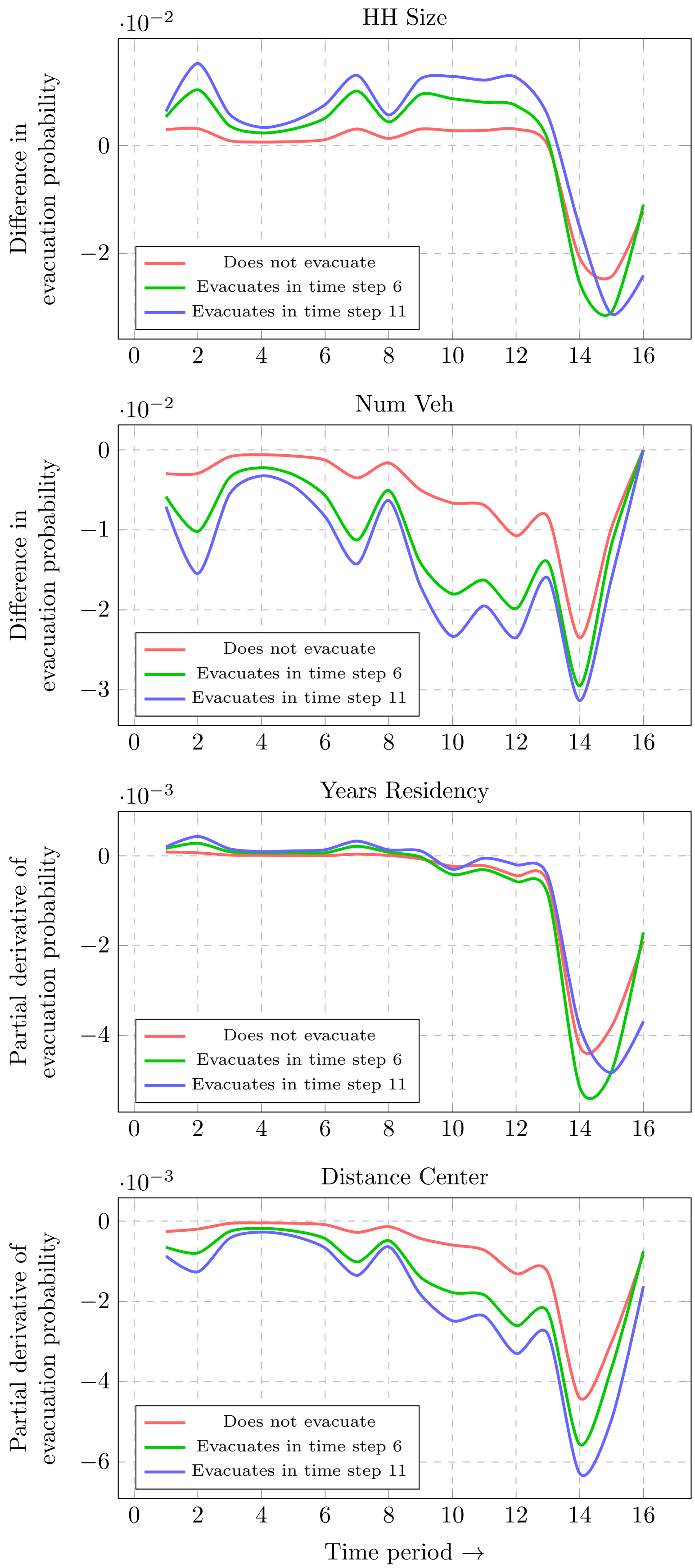

Figure 7: Sensitivity of evacuation probabilities for the model with forecasts.

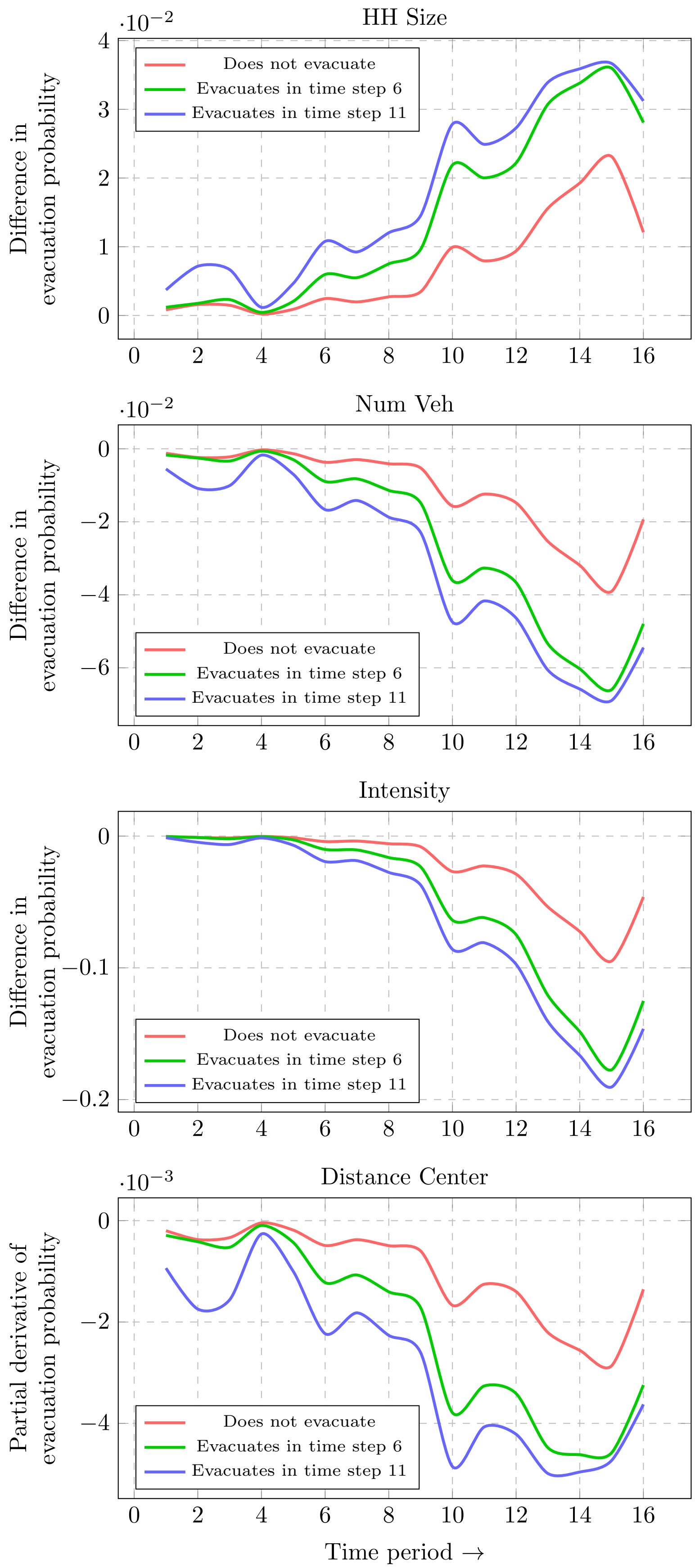

Figure 8: Sensitivity of evacuation probabilities for a logit model.

Using these expressions, we can also plot time-varying sensitivities from the static model as shown in Figure 8. Notice from the figure that the model can distinguish a household that shelters-in-place from those that evacuate in different time periods based on the relative sensitivities. Since the covariates change over time, the scaling term causes the marginals to change with time. For variables that do not have interaction terms such as household size and number of vehicles, the shapes of the marginals are similar and may just differ in sign. For variables that are interacted with NETP, the signs depend on the values of the parameters. For example, consider the distance to the center variable. The marginals can, in theory, switch signs if $\beta_3 + \beta_6\text{NETP}$ is negative for some values of NETP and positive for others. However, for the current data and model specification, we did not observe such phenomena.

In both static and dynamic models, we believe that it not possible to separate the effect of time and the effect of changing hurricane attributes on the choice probabilities. Note that for static models with interaction terms, the interpretation of signs is also not trivial, and equal care must be exercised in inferring the effect of a variable on the choice probabilities. Comparing Figures 7 and 8, the trends in sensitivities are similar except for household size since the static model does not have interaction terms with household size. The sensitivities for the dynamic model with forecasts have relatively higher magnitudes and exhibit sharper changes in choice probabilities over time, probably because of the complex ways in which forecast information influences these calculations.

A major advantage of a static model is that it is easily estimable using off-the-shelf statistical libraries. However, arriving at a specification that provides a good fit and captures complicated time-varying effects would require testing several interaction effects and potentially non-linear transformations of the variables. A related issue is that of modeling forecasts. Forecasted conditions of a storm over multiple future time periods can significantly influence the timing of evacuations. While the dynamic model with forecasts includes them in non-stationary transition probabilities, a static model would require them to be introduced as additional covariates. This increases the number of parameters to be estimated and makes interpretation harder. Also, static models cannot discover latent beliefs on the evolution of uncertainty, as illustrated in the discussion on dynamic models without forecasts.

It appears that the static specifications with time variables can be viewed as a function approximation to the conditional value functions of the dynamic model with forecasts. Linear architectures with polynomial basis functions are widely used in approximate dynamic programming for simplifying value functions, but the approximation quality depends heavily on the set of features used. Using exact results from the dynamic models to engineer features for static frameworks could be a worthwhile topic for future research.

5.4 Validation

Econometric models help understand behavior by identifying the effects of various causal factors on the choices of individuals. These models can also be used for predicting the choices of a household for a completely different hurricane scenario. Further, aggregate-level predictions of demand can also be made, particularly for supply-side models that use data on departure time choices.

Since the dataset used in this paper was small, we tested the predictive abilities of the proposed models using a leave-one-out-cross-validation (LOOCV) approach. In this method, one household is first left out from the data, and the model parameters are estimated from the remaining 276 households. The estimated parameters are then used to find the probability of observing the sequence of actions taken by the left-out household. The validation results for the partial information model without and with forecasts are shown in Figures 9 and 10 respectively.

The x-axis in these plots represents the ID of the household that was left out in the LOOCV procedure. The data points are organized such that the households shown in blue evacuated, while those indicated using the red bars sheltered in place and did not evacuate. Also, the households that evacuated are arranged in increasing order of their departure time as we move from left to right, i.e., those that depart early are to the left of the figure.

Each bar in the top panels of Figures 9 and 10 indicates the probability of observing the exact departure time choice of the left-out household. For each left-out household, we also used the model to predict the probability of evacuating in an 18-hour time window surrounding the actual departure time, except for the households that stayed. These results are shown in the bottom panels of Figures 9 and 10. That is, if a household evacuated in time period 4, the probability with which it could evacuate in time periods 3, 4, or 5 was calculated. Likewise, if a household evacuated in time period 1, the estimated probabilities include evacuating in time periods 1, 2, or 3. For households that stayed, the red bars in the top and bottom panels are identical.

The model performs reasonably well in predicting the actions of households that shelter-in-place but not for those

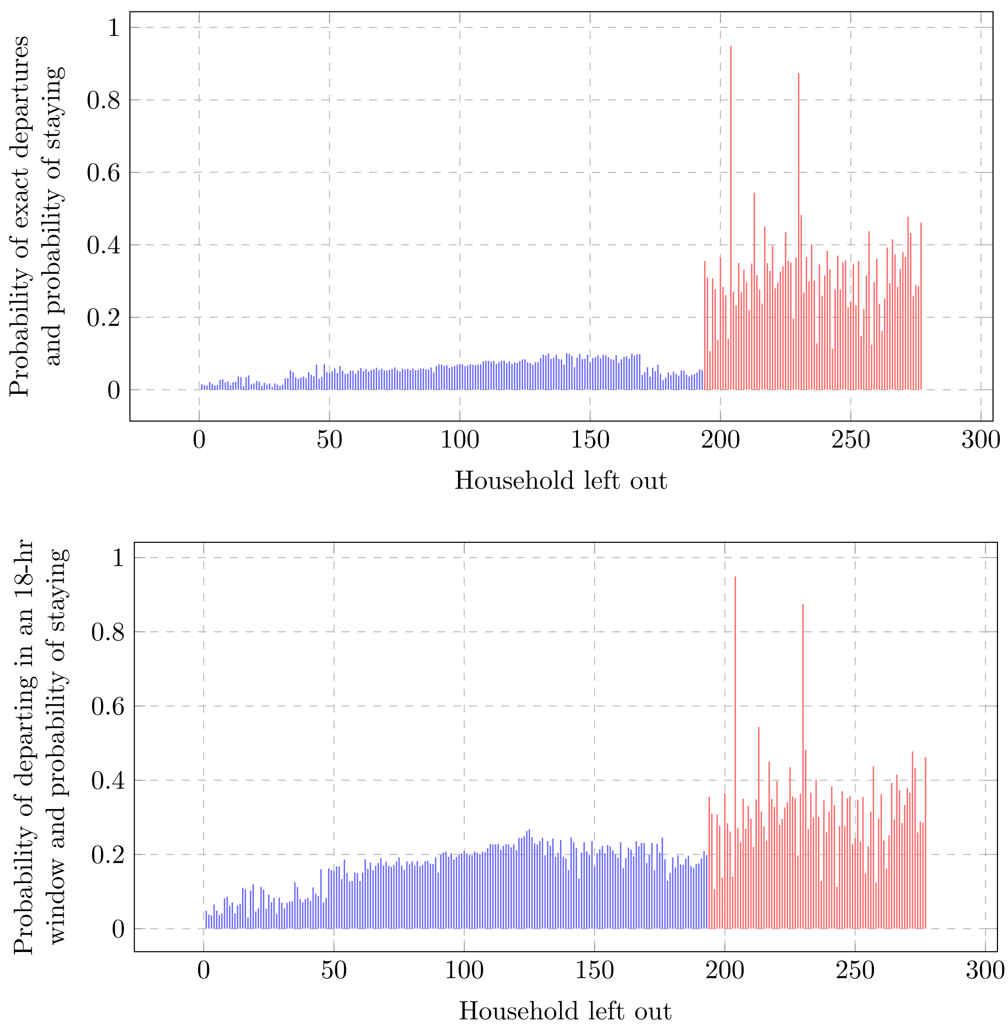

Figure 9: LOOCV for the partial information model without forecasts.

that evacuate, particularly in the early time periods. The prediction probabilities marginally improved when we computed them after considering adjacent time periods. We suspect that the quality of results is due to the presence of a large number of time periods and a sparse set of observations in the beginning, as indicated by the histogram shown in green in Figure 11. We also performed similar experiments after discarding observations from early time periods and noticed an improvement in the predicted probabilities. Between the model with and without forecasts, the one with forecasts was found to have better predictive power. While it is no surprise that the predictions for the 18-hr time windows are better, we have documented them to primarily illustrate the levels of accuracy that can be expected from dynamic choice models for different time granularities. These results can, in turn, provide insights for formulating dynamic network loading models that estimate roadway congestion. We also performed similar validation tests on the static model described in (41), and the results were comparable to the dynamic versions.

Although the validation results at an individual household level appear ordinary, the expected number of predicted departures in each time period (and the predicted number of households that sheltered in place) shown by the blue and red bars in Figure 11 were found to closely match the observed share of departures. To obtain these aggregate numbers, we first predicted the probability of departing in each of the 16 time periods (and the probability of staying back) for a left-out household in the LOOCV procedure. This operation was then repeated 277 times by leaving out one household at a time. Finally, the predicted probabilities were summed across all the left-out households to estimate the expected number of departures.

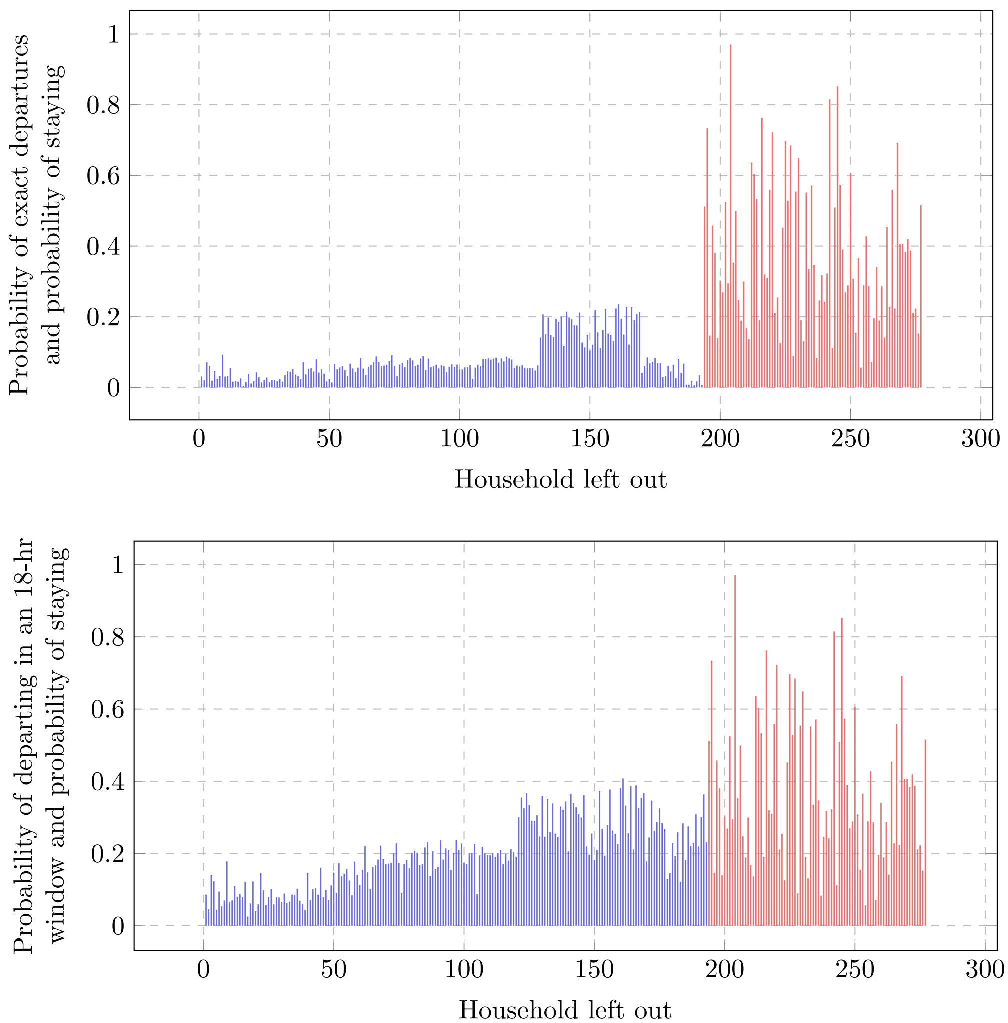

Figure 10: LOOCV for the partial information model with forecasts.

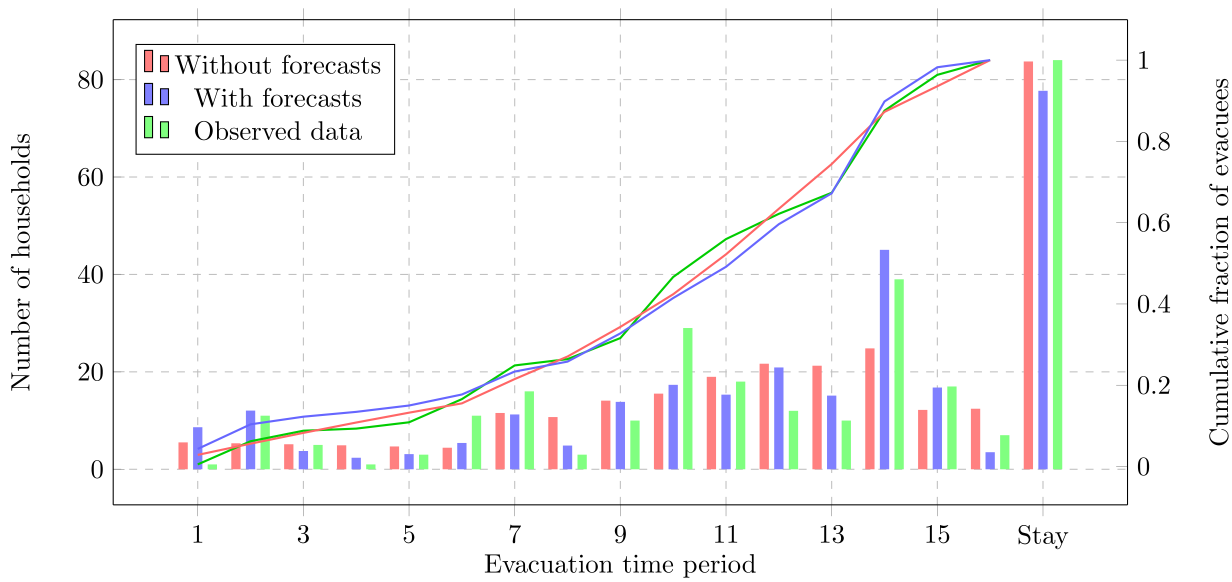

Figure 11: Number of expected evacuations in different time periods.

For each left-out household in the LOOCV procedure, we also estimated the probability of leaving, i.e., the probability of evacuating in any one of the 16 time periods. We computed the average of these values across all left-out households, which evacuated on the same day (or never evacuated). The results are summarized in Table 7. The differences in the average evacuation probabilities were statistically significant in some cases and insignificant in others. However, on average, those who evacuated earlier appear to have a higher predicted probability of not sheltering in place.

Table 7: Average evacuation probabilities.

Evacuation Time Period	*No. of Households*	*Without Forecasts*	*With Forecasts*
Day 1 (1-4)	18	0.734	0.803
Day 2 (5-8)	33	0.713	0.770
Day 3 (9-12)	69	0.709	0.758
Day 4 (13-16)	73	0.696	0.751
Stay	84	0.676	0.623

It is important to note that the validation procedure discussed so far assumes the knowledge of the time-varying hurricane characteristics and forecasts for predicting the choice probabilities across time. In practice, when transferring this model to a different setting, choice predictions can be made one step at a time using the current hurricane characteristics (while still considering an expectation of the utilities from the future), and as the storm unfolds, the evacuation probabilities for the next time period can be re-calculated.

6 Conclusions

In this paper, evacuation-related decisions of households during hurricanes are modeled using an econometric framework to understand the influence of socio-demographic variables and time-varying storm characteristics. Our main contribution is the use of a dynamic discrete choice model in which individuals are assumed to either evacuate or defer their decisions over multiple time periods. If an individual evacuates, the choice process terminates, but if he/she chooses to wait, similar choices are presented in the next time step. These choices are assumed to be made by factoring in current as well as expected future disutilities. The models can also include discount factors that prioritize current disutilities over those experienced in the future. To compute these future disutilities, we assume that households either have some beliefs on the evolution of the storm (which are captured by parameterized transition probability matrices) or utilize forecasts made by the National Hurricane Center. These features were demonstrated using a case study of 277 households that witnessed Hurricane Gustav.

In the future, estimating such models on larger data sets can help predict time-dependent departure rates and address the issues surrounding transferability, which is a critical component in disaster management. The methods used in this paper can also be extended in a few directions. First, the discount factors and belief parameters can also be parameterized as a function of the covariates, and this may capture time-dependent effects and heterogeneity in perceptions and the forward-looking nature of individuals. Model specifications for the dynamic case can also be modified to account for interaction effects. Second, additional data, particularly on roadway travel times, can be collected through surveys or secondary sources and be used to improve the specification of the terminal utilities of the evacuate option. Third, different individuals may begin their choice-making process at different time steps. Having information on when they began thinking about evacuating can enhance model quality. Finally, the optimal stopping problem in this paper is based on expected utility theory, and individuals' choices are assumed to be governed by current and anticipated hurricane scenarios. However, in such dynamic settings, past choices and realized state variables often induce a certain degree of regret and can influence future decision making (Strack and Viefers, 2015). For instance, an individual who waits too long after miscalculating the severity of a hurricane might react in a manner that is inconsistent with expected utility theory and Markovian transition functions. Studies in this direction are limited, and incorporating state and choice history in regret-based objectives could further improve our understanding of evacuation behavior.

Acknowledgments

The authors thank the National Science Foundation (Grant Nos. 0826832 and CMMI-1331269) for financially supporting this research. The statements, findings, and conclusions are those of the authors and do not necessarily reflect the views of the National Science Foundation. Special thanks to Chester Wilmot and Ravindra Gudishala for sharing the household survey data used in the paper. The first author would also like to thank Abdul Pinjari for useful discussions on choice modeling. The authors also appreciate constructive comments made by the reviewers and the associate editor, which have significantly improved our paper.

References

Aguirregabiria, V. and P. Mira (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics 156*(1), 38–67.

Arcidiacono, P. and P. B. Ellickson (2011). Practical methods for estimation of dynamic discrete choice models. *Annu. Rev. Econ. 3*(1), 363–394.

Baker, E. J. (1991). Hurricane evacuation behavior. *International Journal of Mass Emergencies and Disasters 9*(2), 287–310.

Berndt, E. R., B. H. Hall, R. E. Hall, and J. A. Hausman (1974). Estimation and inference in nonlinear structural models. In *Annals of Economic and Social Measurement, Volume 3, number 4*, pp. 653–665. NBER.

Beven, J. and T. Kimberlain (2009). Tropical cyclone report hurricane gustav: 25 august–4 september 2008. *NOAA/National Hurricane Center*.

Cirillo, C., R. Xu, and F. Bastin (2015). A dynamic formulation for car ownership modeling. *Transportation Science 50*(1), 322–335.

Czajkowski, J. (2011). Is it time to go yet? understanding household hurricane evacuation decisions from a dynamic perspective. *Natural Hazards Review 12*(2), 72–84.

Dash, N. and H. Gladwin (2007). Evacuation decision making and behavioral responses: Individual and household. *Natural Hazards Review 8*(3), 69–77.

Dixit, V., T. Montz, and B. Wolshon (2011). Validation techniques for region-level microscopic mass evacuation traffic simulations. *Transportation Research Record: Journal of the Transportation Research Board* (2229), 66–74.

Dixit, V., C. Wilmot, and B. Wolshon (2012). Modeling risk attitudes in evacuation departure choices. *Transportation Research Record 2312*(1), 159–163.

Dow, K. and S. L. Cutter (1998). Crying wolf: Repeat responses to hurricane evacuation orders. *Coastal Management 26*(4), 237–252.

FEMA (1999). *Northwest Florida Hurricane Evacuation Study: Technical Data Report : for Escambia, Santa Rosa, Okaloosa, Walton, Bay, Holmes, Jackson, and Washington Counties, Florida.* Florida State Emergency Management Office.

Fu, H. and C. Wilmot (2004). Sequential logit dynamic travel demand model for hurricane evacuation. *Transportation Research Record: Journal of the Transportation Research Board* (1882), 19–26.

Fu, H., C. Wilmot, H. Zhang, and E. Baker (2007). Modeling the hurricane evacuation response curve. *Transportation Research Record: Journal of the Transportation Research Board* (2022), 94–102.

Glerum, A., O. B. Västberg, E. Frejinger, A. Karlström, M. B. Hugosson, and M. Bierlaire (2015). A dynamic discrete-continuous choice model of car ownership, usage and fuel type. Technical report, Centre Interuniversitaire de Recherche sur les Réseaux d'Entreprise, la Logistique et le Transport.

Gottumukkala, R., R. Kolluru, and M. Smith (2011). Evacuation behavior survey report. Technical report, University of Louisiana at Lafayette.

Gowrisankaran, G. and M. Rysman (2012). Dynamics of consumer demand for new durable goods. *Journal of political Economy 120*(6), 1173–1219.

Gudishala, R. and C. Wilmot (2012). Comparison of time-dependent sequential logit and nested logit for modeling hurricane evacuation demand. *Transportation Research Record: Journal of the Transportation Research Board* (2312), 134–140.

Hasan, S., S. Ukkusuri, H. Gladwin, and P. Murray-Tuite (2011). Behavioral model to understand household-level hurricane evacuation decision making. *Journal of Transportation Engineering 137*(5), 341–348.

Heckman, J. J. (1981). Heterogeneity and state dependence. In *Studies in labor markets*, pp. 91–140. University of Chicago Press.

Hess, S. and D. Palma (2019). Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. *Journal of Choice Modelling 32*, 100170.

Hicks, M. J. and M. L. Burton (2017). Hurricane Harvey: Preliminary estimates of commercial and public sector damages on the Houston metropolitan area. Technical report, Marshall University: Center for Business and Economic Research.

Hotz, V. J. and R. A. Miller (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies 60*(3), 497–529.

Joensen, J. S. (2009). Academic and labor market success: The impact of student employment, abilities, and preferences. Working paper.

Knabb, R. D., D. P. Brown, and J. R. Rhome (2006). Tropical cyclone report, Hurricane Rita, 18-26 september 2005. Technical report, National Hurricane Center.

Knabb, R. D., J. R. Rhome, and D. P. Brown (2011). Tropical cyclone report: Hurricane Katrina, August 23-30, 2005. Technical report, National Hurricane Center.

Li, A. C., N. Xu, L. Nozick, and R. Davidson (2011). Bilevel optimization for integrated shelter location analysis and transportation planning for hurricane events. *Journal of Infrastructure Systems 17*(4), 184–192.

Lindell, M. K. (2008). EMBLEM2: An empirically based large scale evacuation time estimate model. *Transportation research part A: policy and practice 42*(1), 140–154.

Lindell, M. K., J. E. Kang, and C. S. Prater (2011). The logistics of household hurricane evacuation. *Natural hazards 58*(3), 1093–1109.

Lindell, M. K., J.-C. Lu, and C. S. Prater (2005). Household decision making and evacuation in response to Hurricane Lili. *Natural Hazards Review 6*(4), 171–179.

Lindell, M. K. and C. S. Prater (2007). Critical behavioral assumptions in evacuation time estimate analysis for private vehicles: Examples from hurricane research and planning. *Journal of Urban Planning and Development 133*(1), 18–29.

Liu, Y. and C. Cirillo (2018). A generalized dynamic discrete choice model for green vehicle adoption. *Transportation Research Part A: Policy and Practice 114*, 288–302.

Magnac, T. and D. Thesmar (2002). Identifying dynamic discrete decision processes. *Econometrica 70*(2), 801–816.

Mai, T., M. Fosgerau, and E. Frejinger (2015). A nested recursive logit model for route choice analysis. *Transportation Research Part B: Methodological 75*, 100–112.

Melnikov, O. (2013). Demand for differentiated durable products: The case of the US computer printer market. *Economic Inquiry 51*(2), 1277–1298.

Mileti, D. S. and J. H. Sorensen (1990). Communication of emergency public warnings: A social science perspective and state-of-the-art assessment. Technical report, Oak Ridge National Lab., TN (USA).

Miller, R. A. (1984). Job matching and occupational choice. *Journal of Political economy 92*(6), 1086–1120.

Murphy, A. (2015). A dynamic model of housing supply. Working paper.

Murray-Tuite, P. and B. Wolshon (2013). Evacuation transportation modeling: An overview of research, development, and practice. *Transportation Research Part C: Emerging Technologies 27*, 25 – 45.

Nair, H. (2007). Intertemporal price discrimination with forward-looking consumers: Application to the us market for console video-games. *Quantitative Marketing and Economics 5*(3), 239–292.

Pakes, A. (1984). Patents as options: Some estimates of the value of holding european patent stocks.

Peacock, W. G., B. H. Morrow, and H. Gladwin (1997). *Hurricane Andrew: Ethnicity, gender, and the sociology of disasters*. Psychology Press.

Pel, A. J., M. C. Bliemer, and S. P. Hoogendoorn (2012). A review on travel behaviour modelling in dynamic traffic simulation models for evacuations. *Transportation 39*(1), 97–123.

Rambha, T., E. Jafari, and S. D. Boyles (2019). Transportation network issues in evacuations. In K. K. Stephens (Ed.), *New Media in Times of Crisis*, Chapter 7, pp. 144–161. Routledge.

Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica: Journal of the Econometric Society*, 999–1033.

Rust, J. (1988). Maximum likelihood estimation of discrete control processes. *SIAM journal on control and optimization 26*(5), 1006–1024.

Rust, J. (1994). Structural estimation of Markov decision processes. *Handbook of econometrics 4*, 3081–3143.

Sadri, A. M., S. V. Ukkusuri, and H. Gladwin (2017). The role of social networks and information sources on hurricane evacuation decision making. *Natural Hazards Review 18*(3), 04017005.

Sarwar, M. T., P. C. Anastasopoulos, S. V. Ukkusuri, P. Murray-Tuite, and F. L. Mannering (2016). A statistical analysis of the dynamics of household hurricane-evacuation decisions. *Transportation*, 1–20.

Serulle, N. U. and C. Cirillo (2017). The optimal time to evacuate: A behavioral dynamic model on Louisiana resident data. *Transportation Research Part B: Methodological*.

Soika, K. (2006). Evacuation planning in Texas before and after Hurricane Rita. *House Research Organization, Texas House of Representatives, Interim News 79*(2), 1–8.

Strack, P. and P. Viefers (2015). Too proud to stop: Regret in dynamic decisions. *Journal of the European Economic Association*.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

Tuydes, H. and A. Ziliaskopoulos (2006). Tabu-based heuristic approach for optimization of network evacuation contraflow. *Transportation Research Record: Journal of the Transportation Research Board* (1964), 157–168.

Vásquez, W. F., T. J. Murray, and P. Mozumder (2015). Understanding hurricane evacuation planning in the Northeastern and Mid-Atlantic United States. *Natural Hazards Review 17*(1), 04015018.

Whitehead, J. C., B. Edwards, M. Van Willigen, J. R. Maiolo, K. Wilson, and K. T. Smith (2000). Heading for higher ground: factors affecting real and hypothetical hurricane evacuation behavior. *Global Environmental Change Part B: Environmental Hazards 2*(4), 133–142.

Widener, M. J., M. W. Horner, and S. S. Metcalf (2013). Simulating the effects of social networks on a population's hurricane evacuation participation. *Journal of geographical systems 15*(2), 193–209.

Wilmot, C. G., R. Gudishala, et al. (2013). Development of a time-dependent hurricane evacuation model for the new orleans area. Technical Report FHWA/LA.12/494, Louisiana Transportation Research Center.

Wilmot, C. G. and B. Mei (2004). Comparison of alternative trip generation models for hurricane evacuation. *Natural hazards review 5*(4), 170–178.

Wolpin, K. I. (1984). An estimable dynamic stochastic model of fertility and child mortality. *Journal of Political economy 92*(5), 852–874.

Wolshon, B. (2001). "One-way-out": contraflow freeway operation for hurricane evacuation. *Natural hazards review 2*(3), 105–112.

Wu, H.-C., M. K. Lindell, and C. S. Prater (2012). Logistics of hurricane evacuation in Hurricanes Katrina and Rita. *Transportation research part F: traffic psychology and behaviour 15*(4), 445–461.

Xie, C., D.-Y. Lin, and S. T. Waller (2010). A dynamic evacuation network optimization problem with lane reversal and crossing elimination strategies. *Transportation research part E: logistics and transportation review 46*(3), 295–316.

Xu, K., R. A. Davidson, L. K. Nozick, T. Wachtendorf, and S. E. DeYoung (2016). Hurricane evacuation demand models with a focus on use for prediction in future events. *Transportation Research Part A: Policy and Practice 87*, 90–101.

Yi, W., L. Nozick, R. Davidson, B. Blanton, and B. Colle (2017). Optimization of the issuance of evacuation orders under evolving hurricane conditions. *Transportation Research Part B: Methodological 95*, 285 – 304.

A Identification

Not all parameters of a dynamic choice model can be uniquely identified. (See Rust (1994) and Magnac and Thesmar (2002) for a broader discussion on dynamic discrete choice models.) For this reason, we normalized the utility of the stay option to zero. This section provides more explanation on this matter. Assume that the utilities of all three options: evacuate, wait, and stay only include intercepts β, ψ, and ϑ respectively. The conditional and ex ante value functions for this specification can be calculated by solving (44) - (49). In the final period T,

$$v_T(\mathbf{x}_T, \texttt{Evacuate}) = \beta \tag{44}$$

$$v_T(\mathbf{x}_T, \texttt{Stay}) = \vartheta \tag{45}$$

$$\overline{V}_T(\mathbf{x}_T) = \gamma + \ln\big(\exp(\beta) + \exp(\vartheta)\big) \tag{46}$$

For all remaining time periods $t = T-1, T-2, \ldots, 1$,

$$v_t(\mathbf{x}_t, \texttt{Evacuate}) = \beta \tag{47}$$

$$v_t(\mathbf{x}_t, \texttt{Wait}) = \psi + \alpha \int \overline{V}_{t+1}(\mathbf{x}_{t+1}) f(\mathbf{x}_{t+1}|\mathbf{x}_t, \texttt{Wait}) d\mathbf{x}_{t+1} \tag{48}$$

$$\overline{V}_t(\mathbf{x}_t) = \gamma + \ln\Big(\exp(v_t(\mathbf{x}_t, \texttt{Evacuate})) + \exp\left(v_t(\mathbf{x}_t, \texttt{Wait})\right)\Big) \tag{49}$$

Now suppose that we subtract a constant c from the utilities of the evacuate and stay options and $c(\alpha-1)$ from the utility of the wait option. The new conditional and ex ante value functions, denoted by v^c and $\overline{V}^c$ respectively, satisfy the following backward recursion equations.

$$v^c_T(\mathbf{x}_T, \texttt{Evacuate}) = \beta - c = v_T(\mathbf{x}_T, \texttt{Evacuate}) - c \tag{50}$$

$$v^c_T(\mathbf{x}_T, \texttt{Stay}) = \vartheta - c = v_T(\mathbf{x}_T, \texttt{Stay}) - c \tag{51}$$

$$\overline{V}^c_T(\mathbf{x}_T) = \gamma + \ln\big(\exp(\beta - c) + \exp(\vartheta - c)\big) = \overline{V}_T(\mathbf{x}_T) - c \tag{52}$$

For $t = T-1, T-2, \ldots, 1$,

$$v^c_t(\mathbf{x}_t, \texttt{Evacuate}) = \beta - c = v_t(\mathbf{x}_t, \texttt{Evacuate}) - c \tag{53}$$

$$v^c_t(\mathbf{x}_t, \texttt{Wait}) = \psi + c(\alpha - 1) + \alpha \int \overline{V}^c_{t+1}(\mathbf{x}_{t+1}) f(\mathbf{x}_{t+1}|\mathbf{x}_t, \texttt{Wait}) d\mathbf{x}_{t+1} = v_t(\mathbf{x}_t, \texttt{Wait}) - c \tag{54}$$

$$\overline{V}^c_t(\mathbf{x}_t) = \gamma + \ln\Big(\exp(v^c_t(\mathbf{x}_t, \texttt{Evacuate})) + \exp(v^c_t(\mathbf{x}_t, \texttt{Wait}))\Big) = \overline{V}_t(\mathbf{x}_t) - c \tag{55}$$

Notice that the conditional value functions in the second case are c units less than that in the first case for all three choices. Hence, the time-dependent choice probabilities remain the same, and so does the objective function. Thus, it is not possible to identify all three intercepts. For similar reasons, it is not possible to estimate coefficients associated with socio-demographic variables such as income and household size for all three choices. Finally, we cannot have a hurricane-specific attribute in the utility of the stay option if we have intercepts in the evacuate and wait choices. This is because the parameter of the hurricane-specific attribute in the stay option can always be normalized and viewed as a constant.

B Derivatives

In this section, we derive the equations used to calculate the score functions. The parameters α and θ are replaced with $h(\alpha)$ and $h(\theta)$ respectively to transform the maximum likelihood estimation problem into an unconstrained optimization model. For brevity, we will not explicitly include $\boldsymbol{\lambda}$ in the arguments of the value functions and the choice probabilities.

B.1 Perfect Information Models

To compute the partial derivatives of the log-likelihood function, we first calculate the partials of the conditional and ex ante value functions for each time step. Note that $\nabla_{\boldsymbol{\beta}} v_{nt}(\mathbf{y}_{nt}, \texttt{Evacuate}) = \mathbf{y}_{nt}$ and that the partial derivative of

$v_{nt}(\mathbf{y}_{nt}, \texttt{Evacuate})$ with respect to the other parameters is 0. The partials of $v_{nT}(\mathbf{y}_{nT}, \texttt{Stay})$ with respect to all the parameters is 0.

Derivatives of the conditional value functions of the wait option are relatively involved since they depend on the ex ante value function (defined in (12)) from the future time step (see (14)). Hence, they have to be computed using a backward pass by recursively solving (56) - (59) for $t = T, T-1, \ldots, 2$. At each time period we first find the derivatives of the ex ante value functions using

$$\nabla_{\boldsymbol{\lambda}} \overline{V}_{nt}(\mathbf{y}_{nt}) = \frac{\sum_{a_t \in A_t} \exp\big(v_{nt}(\mathbf{y}_{nt}, a_t)\big) \nabla_{\boldsymbol{\lambda}} v_{nt}(\mathbf{y}_{nt}, a_t)}{\sum_{a_t \in A_t} \exp\big(v_{nt}(\mathbf{y}_{nt}, a_t)\big)} \tag{56}$$

The partials of the conditional value functions of the wait option are then updated as follows

$$\nabla_{\boldsymbol{\beta}} v_{n,t-1}(\mathbf{y}_{n,t-1}, \texttt{Wait}) = h(\alpha) \nabla_{\boldsymbol{\beta}} \overline{V}_{nt}(\mathbf{y}_{nt}) \tag{57}$$

$$\nabla_{\boldsymbol{\psi}} v_{n,t-1}(\mathbf{y}_{n,t-1}, \texttt{Wait}) = \mathbf{y}_{n,t-1} + h(\alpha) \nabla_{\boldsymbol{\psi}} \overline{V}_{nt}(\mathbf{y}_{nt}) \tag{58}$$

$$\frac{\partial}{\partial \alpha} v_{n,t-1}(\mathbf{y}_{n,t-1}, \texttt{Wait}) = h'(\alpha) \overline{V}_{nt}(\mathbf{y}_{nt}) + h(\alpha) \frac{\partial}{\partial \alpha} \overline{V}_{nt}(\mathbf{y}_{nt}) \tag{59}$$

Using the derivatives of the conditional value functions, one can estimate the derivatives of the time-dependent choice probabilities. For time periods $T-1, T-2, \ldots, 1$, we may write

$$\nabla_{\boldsymbol{\lambda}} \mathbb{P}[\texttt{Wait}|\mathbf{y}_{nt}] = \frac{-\exp\big(v_{nt}(\mathbf{y}_{nt}, \texttt{Evacuate}) - v_{nt}(\mathbf{y}_{nt}, \texttt{Wait})\big)}{\Big(1 + \exp\big(v_{nt}(\mathbf{y}_{nt}, \texttt{Evacuate}) - v_{nt}(\mathbf{y}_{nt}, \texttt{Wait})\big)\Big)^2} \Big(\nabla_{\boldsymbol{\lambda}} v_{nt}(\mathbf{y}_{nt}, \texttt{Evacuate}) - \nabla_{\boldsymbol{\lambda}} v_{nt}(\mathbf{y}_{nt}, \texttt{Wait})\Big) \tag{60}$$

Observe that $\nabla_{\boldsymbol{\lambda}} \mathbb{P}[\texttt{Evacuate}|\mathbf{y}_{nt}]$ is simply $\mathbf{1} - \nabla_{\boldsymbol{\lambda}} \mathbb{P}[\texttt{Wait}|\mathbf{y}_{nt}]$. Similar expressions can be derived for time period T by replacing the wait option in (60) with stay.

B.2 Partial Information Models

B.2.1 Without Forecasts

The score vector for the partial information model without forecasts is calculated in a similar way as the perfect information case except that the derivatives must be computed for every state and must take into account the effects of the belief parameter. First, note that for a household n and time step t, $\nabla_{\boldsymbol{\beta}} v_{nt}(\mathbf{x}_{nt}, \texttt{Evacuate}) = \mathbf{x}_{nt}$ for all $\mathbf{x}_{nt} \in X_{nt}$ and the partial derivatives of $v_{nT}(\mathbf{x}_{nT}, \texttt{Stay})$ with respect to all parameters is 0. The derivatives of the conditional value functions on the other hand are calculated from the following backward recursion steps. For $t = T, T-1, \ldots, 2$,

$$\nabla_{\boldsymbol{\lambda}} \overline{V}_{nt}(\mathbf{x}_{nt}) = \frac{\sum_{a_t \in A_t} \exp\big(v_{nt}(\mathbf{x}_{nt}, a_t)\big) \nabla_{\boldsymbol{\lambda}} v_{nt}(\mathbf{x}_{nt}, a_t)}{\sum_{a_t \in A_t} \exp\big(v_{nt}(\mathbf{x}_{nt}, a_t)\big)} \, \forall \mathbf{x}_{nt} \in X_{nt} \tag{61}$$

For all $\mathbf{x}_{n,t-1} \in X_{n,t-1}$, the partials of the conditional value functions of the wait option satisfy

$$\nabla_{\boldsymbol{\beta}} v_{n,t-1}(\mathbf{x}_{n,t-1}, \texttt{Wait}) = h(\alpha) \sum_{\mathbf{x}_{nt} \in X_{nt}} f\big(\mathbf{x}_{nt}|\mathbf{x}_{n,t-1}; h(\theta)\big) \nabla_{\boldsymbol{\beta}} \overline{V}_{nt}(\mathbf{x}_{nt}) \tag{62}$$

$$\nabla_{\boldsymbol{\psi}} v_{n,t-1}(\mathbf{x}_{n,t-1}, \texttt{Wait}) = \mathbf{x}_{n,t-1} + h(\alpha) \sum_{\mathbf{x}_{nt} \in X_{nt}} f\big(\mathbf{x}_{nt}|\mathbf{x}_{n,t-1}; h(\theta)\big) \nabla_{\boldsymbol{\psi}} \overline{V}_{nt}(\mathbf{x}_{nt}) \tag{63}$$

$$\frac{\partial}{\partial \alpha} v_{n,t-1}(\mathbf{x}_{n,t-1}, \texttt{Wait}) = \sum_{\mathbf{x}_{nt} \in X_{nt}} f\big(\mathbf{x}_{nt}|\mathbf{x}_{n,t-1}; h(\theta)\big) \Big\{ h'(\alpha) \overline{V}_{nt}(\mathbf{x}_{nt}) + h(\alpha) \frac{\partial}{\partial \alpha} \overline{V}_{nt}(\mathbf{x}_{nt}) \Big\} \tag{64}$$

$$\frac{\partial}{\partial \theta} v_{n,t-1}(\mathbf{x}_{n,t-1}, \texttt{Wait}) = h(\alpha) \sum_{\mathbf{x}_{nt} \in X_{nt}} \Big\{ f'\big(\mathbf{x}_{nt}|\mathbf{x}_{n,t-1}; h(\theta)\big) \overline{V}_{nt}(\mathbf{x}_{nt}) + f\big(\mathbf{x}_{nt}|\mathbf{x}_{n,t-1}; h(\theta)\big) \frac{\partial}{\partial \theta} \overline{V}_{nt}(\mathbf{x}_{nt}) \Big\} \tag{65}$$

Derivatives of the choice probabilities are derived using (60), which are then used to calculate the score functions as defined in (30).

B.2.2 With forecasts

Since α is constrained, we transform it into an unconstrained variable using $h(\alpha)$. Equations (66)-(68) must be solved using a backward pass to obtain the partial derivatives of the ex ante value function. The partial derivatives of the conditional value functions and the choice probabilities can then be easily estimated by differentiating (31), (32), and by using (60).

$$\nabla_{\boldsymbol{\beta}}\overline{V}_{ntt'}(\mathbf{x}_{nt'}) = \Bigg\{ \exp(\boldsymbol{\beta}^\mathsf{T}\mathbf{x}_{nt'})\mathbf{x}_{nt'} + \exp\Big(\boldsymbol{\psi}^\mathsf{T}\mathbf{x}_{nt'} + h(\alpha)\sum_{\substack{\mathbf{x}_{n,t'+1}\in\\ X_{n,t'+1}}} f_{tt'}(\mathbf{x}_{n,t'+1})\overline{V}_{nt,t'+1}(\mathbf{x}_{n,t'+1})\Big) \times \Big(h(\alpha)\sum_{\substack{\mathbf{x}_{n,t'+1}\in\\ X_{n,t'+1}}} f_{tt'}(\mathbf{x}_{n,t'+1})\nabla_{\boldsymbol{\beta}}\overline{V}_{nt,t'+1}(\mathbf{x}_{n,t'+1})\Big)\Bigg\}\frac{1}{\exp\big(\overline{V}_{ntt'}(\mathbf{x}_{nt'})-\gamma\big)} \tag{66}$$

$$\nabla_{\boldsymbol{\psi}}\overline{V}_{ntt'}(\mathbf{x}_{nt'}) = \Bigg\{ \exp\Big(\boldsymbol{\psi}^\mathsf{T}\mathbf{x}_{nt'} + h(\alpha)\sum_{\substack{\mathbf{x}_{n,t'+1}\in\\ X_{n,t'+1}}} f_{tt'}(\mathbf{x}_{n,t'+1})\overline{V}_{nt,t'+1}(\mathbf{x}_{n,t'+1})\Big) \times \Big(\mathbf{x}_{nt'} + h(\alpha)\sum_{\substack{\mathbf{x}_{n,t'+1}\in\\ X_{n,t'+1}}} f_{tt'}(\mathbf{x}_{n,t'+1})\nabla_{\boldsymbol{\psi}}\overline{V}_{nt,t'+1}(\mathbf{x}_{n,t'+1})\Big)\Bigg\}\frac{1}{\exp\big(\overline{V}_{ntt'}(\mathbf{x}_{nt'})-\gamma\big)} \tag{67}$$

$$\frac{\partial}{\partial\alpha}\overline{V}_{ntt'}(\mathbf{x}_{nt'}) = \frac{1}{\exp\big(\overline{V}_{ntt'}(\mathbf{x}_{nt'})-\gamma\big)}\Bigg\{ \exp\Big(\boldsymbol{\psi}^\mathsf{T}\mathbf{x}_{nt'} + h(\alpha)\sum_{\substack{\mathbf{x}_{n,t'+1}\in\\ X_{n,t'+1}}} f_{tt'}(\mathbf{x}_{n,t'+1})\overline{V}_{nt,t'+1}(\mathbf{x}_{n,t'+1})\Big) \times \Big(h(\alpha)\sum_{\substack{\mathbf{x}_{n,t'+1}\in\\ X_{n,t'+1}}} f_{tt'}(\mathbf{x}_{n,t'+1})\frac{\partial}{\partial\alpha}\overline{V}_{nt,t'+1}(\mathbf{x}_{n,t'+1}) + h'(\alpha)\sum_{\substack{\mathbf{x}_{n,t'+1}\in\\ X_{n,t'+1}}} f_{tt'}(\mathbf{x}_{n,t'+1})\overline{V}_{nt,t'+1}(\mathbf{x}_{n,t'+1})\Big)\Bigg\} \tag{68}$$

C Value Function Approximations

For the data considered used in this paper, dynamic programs could be accurately solved using backward recursion. When dealing with larger state spaces or dynamic-continuous covariates, a simulation-based architecture can be used to approximate value functions. Serulle and Cirillo (2017), for instance, assume that individuals calculate the expectations in the Bellman equations using data from a limited number of time periods (one in their paper) and set the expected future utilities from later periods to zero. This assumption was motivated on the basis of perception limitations of humans in sequential-decision making settings. However, it implicitly presupposes myopic behavior. The discount factors in our experiments are closer to 1, indicating that individuals are more forward-looking than anticipated.

Take, for instance, the perfect information model introduced in Section 4.1. For the last two time periods, an approximate value function that uses the utility from one future time period and the exact method result in the same choice probabilities. Consider the time period $T-2$. The conditional value function for a decision-maker, using the exact method, is

$$\boldsymbol{\psi}^\mathsf{T}\mathbf{y}_{T-2} + \alpha\gamma + \alpha\ln\Big\{\exp(\boldsymbol{\beta}^\mathsf{T}\mathbf{y}_{T-1}) + \exp\Big(\boldsymbol{\psi}^\mathsf{T}\mathbf{y}_{T-1} + \underbrace{\alpha\gamma + \alpha\ln\big\{1+\exp(\boldsymbol{\beta}^\mathsf{T}\mathbf{y}_T)\big\}}_{\text{Utility from last time period}}\Big)\Big\} \tag{69}$$

If expected utilities beyond the immediate future time period are ignored, the highlighted portion in (69) is set to 0. For $T-3$, two such utility expressions are set to 0 and so on. Hence, the approximation errors get compounded as we iterate backward. To illustrate these differences, we estimated a perfect information model using the exact method with a single covariate—Distance Center. The evacuate and wait intercepts were found to be 3.534 and −4.452, the coefficients of the distance variable were −0.014 and 0.002, and the discount factor was 1. All four utility parameters were statistically significant. A plot of the evacuation probabilities over time, calculated using the exact and the approximate approach, for a household that evacuated in time period 11 is shown in Figure 12.

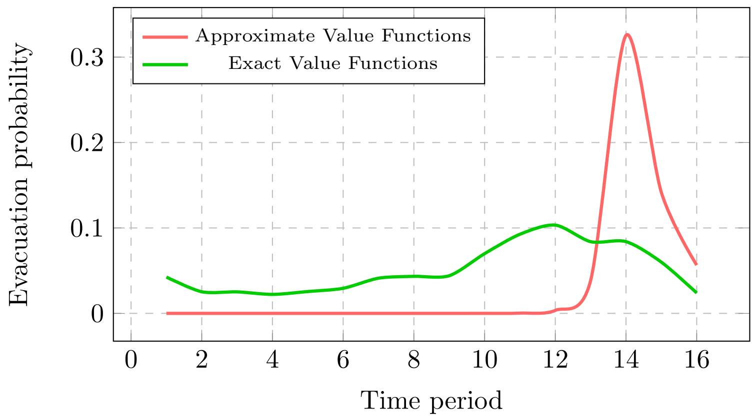

Figure 12: Estimated evacuation probabilities over time for a household that evacuated in time period 11.

The distance between the household and the center of the hurricane drops monotonically from 1184 km in time step 11 to 403 km in time step 16. The predicted probability of evacuation using the approximate value functions is nearly zero at time step 11 and shows a delayed response to decreasing distance values since it only uses information from one future time step. Similar trends were noticed for other households.

We used the perfect information model to showcase these differences only because the expressions for the conditional value functions are straightforward. For partial information models, one can imagine that forecasted conditions several time periods into the future can trigger early evacuations depending on the level of severity of the hurricane. Approximate value functions will, however, be poor at predicting such behavior because of their non-anticipatory nature. Thus, approximations of the ex ante value functions must be well-designed using simulation to avoid significant differences in the conditional value functions and choice probabilities, and consequently, in the parameter estimates.