A Relational Event Model for Social Action, with Application to the World Trade Center Disaster[*][†]

Carter T. Butts[‡]

May 12, 2006

This is a draft document. Please do not distribute or cite without permission.

Abstract

Interpersonal interaction over short time scales is frequently understood in terms of actions, which can be thought of as discrete events in which one individual emits a behavior directed at one or more other entities in his or her environment (possibly including him or herself). Here, we introduce a highly flexible framework for modeling actions within social settings, which permits likelihood-based inference for behavioral mechanisms with complex dependence. The utility of the framework is illustrated via an application to dynamic modeling of responder radio communications during the early hours of the World Trade Center disaster.

Keywords: social action, hazard modeling, dynamic networks, crisis settings, World Trade Center

1 Introduction

Human activity over short time scales is frequently understood in terms of *actions*, which can be thought of as discrete events in which one individual emits a behavior directed at one or more other entities in his or her environment (possibly including him or herself). Actions are often assumed to have social meaning for the actor and/or outside observers (Goffman, 1963),

[*]This work was supported by NIH award 5 R01 DA012831-05 and NSF ITR award IIS-0331707.

[†]The author would like to thank David Mendonça, Miruna Petrescu-Prahova, Ryan Acton, and Ben Lind for their helpful input.

[‡]Department of Sociology and Institute for Mathematical Behavioral Sciences, University of California, Irvine; SSPA 2145, Irvine, CA 92697; buttsc@uci.edu

and have been posited to result from processes ranging from learning (Macy, 1991) and imitation (Miller et al., 2001) to optimization (Coleman, 1990); while many of these perspectives are in substantial disagreement about the mechanisms which underlie action, they arguably have a mutual intersection around a common, "core" notion of action per se. This suggests, from a modeling perspective, that a framework centered on a minimal conception of social action – but with the capacity to incorporate various competing theoretical "overlays" – is likely to prove especially fruitful. To be useful, such a framework must also permit inference from behavioral data, so as to allow for the estimation of the relative strengths of potential mechanisms as well as for principled selection among competing models. Here, we introduce such a framework, organized around the notion of *relational events*. This framework allows for the modeling of complex dependence among actions, differential treatment of action by type, and influences due to exogenous covariates. The framework also supports likelihood-based inference, thereby facilitating the empirical evaluation of competing explanations for social action within particular settings.

One context in which such a framework is particularly relevant is that of behavior during crisis settings. Actors operating within such settings face severe environmental constraints, as well as cognitive limitations (e.g., narrowed attention) and an unstable social context (e.g., due to disrupted role performance (Dynes, 1970, p.176)). Nevertheless, both trained and untrained actors will attempt to respond to hazardous conditions, taking action to obtain safety for themselves and others (Quarantelli, 1960; Mileti et al., 1975; Abe, 1976; Noji, 1997). While many factors affecting behavior in crisis settings have been identified through past field studes (see e.g. Drabek, 1986, for a review), sorting through them has proven difficult: without a systematic way to combine and compare mechanisms, it is rarely possible to adjudicate among competing explanations (much less to make quantitative statements of relative importance). Using the modeling framework developed here, such comparisons can be performed using data (such as radio transcripts or event logs) which can often be obtained in field settings. Here, we illustrate this potential via the application of the relational event framework to radio communication data from the World Trade Center (WTC) disaster.

2 Relational Event Model

As we have emphasized, a more complete understanding of behavior in settings such as crisis situations requires the integration of cognitive, behavioral, and social/contextual processes. To that end, we provide here a framework for the dynamic modeling of social action, incorporating all three factors. Unlike traditional agent based modeling schemes, our approach draws upon event history analysis (e.g., Blossfeld and Rohwer, 1995) to formulate models which can be fit directly to data. On the other hand, our approach is also unlike that employed in most familiar statistical contexts, in that we allow for complex structures of historical dependence among observed events.[1] By building models which are both theoretically informed and inferentially tractable, we hope to obtain important new insights into crisis behavior.

The central element of our modeling approach is the *relational event*, or *action*, which is defined as a discrete event generated by a social actor (the "sender") and directed towards one or more targets (the "receivers," who may or may not be actors themselves); here we restrict ourselves to the single-receiver case. We represent actions by tuples of the form $a = (i, j, k, t)$, where $i \in \mathcal{S}$ represents the *sender* of the action, $j \in \mathcal{R}$ represents the *receiver* of the action, $k \in \mathcal{C}$ represents the *action type*, and $t \in \mathbb{R}$ represents the *time* at which the action is taken. For purposes of the present development, we will assume that each action is associated with a single time point. For convenience, we also define functions s, r, c, and τ, which return (for any given action) the sender, receiver, action type, and time (respectively).

Given an ordered set of actions $a_1, a_2, \ldots$, let the set $A_t = \left\{a_i : \tau(a_i) \leq t\right\}$ consist of all actions taken on or before time t. For convenience, we also define a *null action*, a_0, such that $\tau(a_0) = 0$, and (without loss of generality) take $\tau(a_i) \geq 0 \; \forall \; a_i \in A_t$. By assumption, $a_0 \notin A_t$. For our current family of models, we assume that actions occur via an inhomogeneous Poisson process such that acts arise independently conditional on the realized history of previous actions (along, possibly, with covariates). If h is the hazard function of this process, and S the associated survival function, then we may write

[1] Though see Snijders (2005) for a closely related family of models.

the likelihood of A_t as

$$p(A_t) = \left[\prod_{i=1}^{|A_t|} \left[h\left(\tau\left(a_i\right) - \tau\left(a_{i-1}\right) \middle| s\left(a_i\right), r\left(a_i\right), c\left(a_i\right), A_{\tau(a_{i-1})}\right) \right.\right.$$
$$\left.\left. \times \left(\prod_{j=1}^{|\mathcal{S}|}\prod_{k=1}^{|\mathcal{R}|}\prod_{l=1}^{|\mathcal{C}|} S\left(\tau\left(a_i\right) - \tau\left(a_{i-1}\right) \middle| j, k, l, A_{\tau(a_{i-1})}\right)\right)\right]\right] \tag{1}$$
$$\times \left[\prod_{i=1}^{|\mathcal{S}|}\prod_{j=1}^{|\mathcal{R}|}\prod_{k=1}^{|\mathcal{C}|} S\left(t - \tau\left(a_{|A_t|}\right) \middle| i, j, k, A_t\right)\right].$$

Intuitively, Equation 1 traces the history of A_t, incorporating both the likelihoods of events which did occur (the elements of A_t) and the likelihoods of the associated "non-events" (actions which could have been taken in each instant, but were not). Given this general form, we may then specify particular sub-families by appropriate selection of h and S. One obvious choice in this regard is to assume that each potential action has a constant hazard of occurrence given a particular prior event history (i.e., a piecewise constant latent hazard model). This amounts to the assumption that the waiting time from one event to the next is conditionally exponentially distributed, and hence we can posit some rate function λ such that $h(t) = \lambda$ and $S(t) = e^{-\lambda t}$. λ, in turn, may be a function of sender, receiver, action type, and past action history, as well as exogenous covariates. While many choices of λ are possible, certain properties do suggest themselves as starting points; these are considered in Section 2.2 below.

For brevity of notation, let us define $s_i = s(a_i)$, $r_i = r(a_i)$, $c_i = c(a_i)$, $\tau_i = \tau(a_i)$, and $\lambda_{ijkl} = \lambda(i, j, k, A_l, X, \theta)$ (where i is a sender, j is a receiver, k is an action type, $0 \leq l \leq t$ is an event time, X is a collection of covariates, and θ is a vector of parameters). Under the piecewise constant hazard model, we may now substitute the implied definitions of h and S into Equation 1, thereby obtaining the likelihood

$$p(A_t) = \left[\prod_{i=1}^{|A_t|} \lambda_{s_i r_i c_i \tau_{i-1}} \prod_{j=1}^{|\mathcal{S}|}\prod_{k=1}^{|\mathcal{R}|}\prod_{l=1}^{|\mathcal{C}|} \exp\left(-\lambda_{jkl\tau_{i-1}}\left(\tau_i - \tau_{i-1}\right)\right)\right]$$
$$\times \left[\prod_{i=1}^{|\mathcal{S}|}\prod_{j=1}^{|\mathcal{R}|}\prod_{k=1}^{|\mathcal{C}|} \exp\left(-\lambda_{jkl\tau_{|A_t|}}\left(t - \tau_{|A_t|}\right)\right)\right]. \tag{2}$$

Where λ incorporates unknown parameters, Equation 2 may be used to estimate them (see Section 2.3 below). This is, however, contingent upon fully observed timing information. For data of the sort employed here, however, this assumption is problematic; before turning to the question of how λ may be parameterized, then, we must first determine how the action model may be adapted to data for which timing information is more limited.

2.1 Ordinal Data Likelihood

Where A_t is fully known, the likelihood of Equation 2 provides an adequate basis for subsequent inference. In general, however, this is not the case – while we may know the *order* in which events occur, we do not always have access to exact inter-event times. This is especially true when working with transcript data (like that employed here), for which the exact timing associated with speech events may not have been recorded. If we assume that τ is known only up to an order-preserving transformation, what can be said regarding the likelihood of A_t? Plainly, Equation 2 cannot be used directly, as the inter-event time intervals $(\tau_i - \tau_{i-1})$ are not invariant under the appropriate class of transformations. Indeed, if events in A_t can only be ordered, it follows that the associated likelihood can only be based on which of the various possible events appears next in the τ-induced sequence. For the piecewise constant hazard model, it happens that just such a result can be obtained. We begin by noting that, under the model of Equation 2, the waiting time for any given event a_i following some event a_{i-1}, conditional on the non-occurrence of all other events, is exponentially distributed with parameter $\lambda_{s_i r_i c_i \tau_{i-1}}$. The probability that a_i is the first of the possible events to occur is equivalent to the probability that the waiting time for a_i is the minimum waiting time for all potential events;[2] thus, under the piecewise constant hazard model, the probability that a_i occurs first is equal to the probability that a random variable $W(s_i, r_i, c_i, \tau_{i-1})$ is equal to $\min\{W(1,1,1,\tau_{i-1}), \ldots, W(|\mathcal{S}|, |\mathcal{R}|, |\mathcal{C}|, \tau_{i-1})\}$ where $W(i,j,k,l) \sim \exp(\lambda_{ijkl})$. This probability, in turn, is obtained via the following theorem:

Theorem 1. *Let $X_1, \ldots, X_n$ be independent, exponentially distributed random variables with rate parameters $\eta_1, \ldots, \eta_n$. Then,* $\Pr\big(x_i = \min\{x_1, \ldots, x_n\}\big) = \eta_i / \sum_{j=1}^n \eta_j$.

Proof. Without loss of generality, let $Y = \{X_j : j \neq i\}$ refer to the X variables other than X_i; for clarity of notation, we will relabel the elements of this set as $Y_1, \ldots, Y_{n-1}$ with associated rate parameters $\eta'_1, \ldots, \eta'_{n-1}$. By

[2]Since we are assuming a continous waiting time distribution, we may ignore the probability-zero case in which two events occur at precisely the same instant.

definition, then, $\Pr\big(x_i = \min\{x_1, \ldots, x_n\}\big) = \Pr\big(X_i < \min\{Y_1, \ldots, Y_{n-1}\}\big)$. From the definition of the exponential density, this gives us

$$\Pr\big(X_i < \min\{Y_1, \ldots, Y_m\}\big) = \int_0^\infty \int_{x_i}^\infty \cdots \int_{x_i}^\infty \eta_i e^{-\eta_i x_i} \left(\prod_{j=1}^{n-1} \eta_i e^{-\eta'_j y_j} \mathrm{d}y_j \right) \mathrm{d}x_i.$$

Note that since the Y variables depend only on the value of X_i, we may safely integrate them out, giving us

$$= \int_0^\infty \eta_i e^{-\eta_i x_i} e^{-\sum_{j=1}^{n-1} \eta'_j x_i} \mathrm{d}x_i.$$

Integrating over the range of X_i, we have

$$= \frac{-\eta_i}{\eta_i + \sum_{j=1}^{n-1} \eta'_j} e^{-\left(\eta_i + \sum_{j=1}^{n-1} \eta'_j\right) x_i} \Bigg|_0^\infty,$$

which reduces simply to

$$= \frac{\eta_i}{\eta_i + \sum_{j=1}^{n-1} \eta'_j}.$$

Since, by construction, $\sum_{j=1}^{n-1} \eta'_j = \sum_{j=1}^{n} \eta_j - \eta_i$, we may rewrite this expression as

$$= \frac{\eta_i}{\sum_{j=1}^{n} \eta_j},$$

which completes the proof.

□

Theorem 1 provides us with an surprisingly simple result: the probability that a particular event a_i will be the next in an event sequence (under the piecewise constant hazard model) is equal to the occurence rate for a_i, divided by the sum of the rates for all possible events which might occur (including a_i itself). Since successive events are conditionally independent, it follows that the likelihood of A_t under temporally ordinal data is merely a product of multinomial likelihoods. Specifically,

$$p(A_t) = \prod_{i=1}^{|A_t|} \left[\frac{\lambda_{s_i r_i c_i \tau_{i-1}}}{\sum_{j=1}^{|\mathcal{S}|} \sum_{j=1}^{|\mathcal{R}|} \sum_{j=1}^{|\mathcal{C}|} \lambda_{jkl\tau_{i-1}}} \right]. \tag{3}$$

Ironically, this expression is even simpler than that of Equation 2. This simplicity is not without cost, however: in addition to the fact that information is lost in the conversion from ratio to ordinal scaling, the particular form of Equation 3 allows λ to be identified only up to a constant factor. In practice, this is not a terribly onerous restriction, since we are generally interested in relative rates rather than absolute pace of interaction. However, it should be noted that this does affect extrapolative simulation, in that the average communication rate cannot be determined without additional information. It should thus be emphasized that exact timing information should be used where available, although the relational event model can be usefully applied to data for which only order is known.

2.2 Construction of the Rate Function

While Equations 2 and 3 provide expressions for the the stochastic component of the relational event model, the dynamic evolution of the event system itself is driven primarily by the rate function, λ. As indicated above, we presume that λ may in general depend upon sender, reciever, action type, past event history, and/or exogenous covariates, in addition to unknown parameters. Substantively, such dependence allows us to accomplish a number of modeling goals. First, it is desirable to be able to incorporate sender and receiver effects, i.e. differential tendencies for certain persons or objects (or persons/objects with certain properties) to send or receive action. Second, it is important to be able to include dyadic covariates (e.g., homohpily or physical proximity) which may impact the chance of interaction between any two individuals. Third, the history of past action should impact future behavior, in accordance with known cognitive and behavioral principles. To capture such phenomena within a flexible, interpretable framework, we parameterize the rate function as

$$\lambda(i, j, k, A_l, X, \theta) = \begin{cases} \exp\left(\lambda_0 + \theta^T u\left(i, j, k, A_l, X\right)\right) & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}, \tag{4}$$

where $i \in 1, \ldots, |\mathcal{S}|$ is a sender index, $j \in 1, \ldots, |\mathcal{R}|$ is a receiver index, $k \in 1, \ldots, |\mathcal{C}|$ is an action type index, l is a time point, X is a set of exogenous covariates, $\theta \in \mathbb{R}^p$ is a parameter vector, $u : (i, j, k, A_l, X) \mapsto \mathbb{R}^p$ is a vector of sufficient statistics, and λ_0 is a "pacing constant." Under the ordinal data model, λ_0 is arbitrary (and hence may be taken to be equal to zero without loss of generality); we separate it from the θ effects for this reason. Intuitively, u indexes the various influences which increase or decrease the relative rates of incidence across events. These effects are weighted by θ, such that each unit change in u_i for an event multiplies its relative log-rate by $\exp(\theta_i)$.

Under the parameterization of Equation 4, construction of the rate function amounts to the choice of sufficient statistics (u) which are to be included. While a wide range of statistics may be contemplated, we here present five basic categories of effects which are motivated by the specific properties of the WTC communication data.

2.2.1 Fixed Effects

Within an emergency setting, it is unlikely that all actors will manifest the same base tendency to participate in communication with others. Such differential participation may reflect unobserved heterogeneity in situational awareness, training, or institutional role, as well as differences in local context. For instance, responders whose location places them in imminent danger are unlikely to spend long periods of time engaged in radio communication, relative to those whose locations afford them a greater degree of safety. To capture the impact of such factors when they cannot be measured directly, we propose to include fixed effects for participation in the relational event system. To parameterize such effects, we add $N = |\mathcal{S}| = |\mathcal{R}|$ statistics of the form $u_m(i, j, k, A_l, X) = I(m \in \{i, j\})$, where I is the standard indicator function. The corresponding θ parameters then represent logged rate multipliers for all events having the corresponding individuals as senders or recievers. (Such parameters fulfill the same role as the expansiveness/popularity parameters of the well-known p_1 model (Holland and Leinhardt, 1981).)

It should be noted that, for the ordinal model (or the exact timing model with λ_0 included), the likelihood will not identify all N fixed effect parameters. This may be easily resolved by fixing one parameter to 0, in which case the others should be interpreted as providing log rate multipliers relative to the reference actor. Other linear constraints (e.g., requiring that the statistics sum to 0) may also be applied, if desired. We employ the former solution for the analyses employed here, treating the first individual in each network as the reference actor.

2.2.2 Persistence

Another basic mechanism which is easily captured through the relational event model is *persistence,* or the tendency of past contacts to become future contacts. In particular, let $d(i, j, A_k)$ represent the accumulated volume of communication from actor i to actor j by time k, and let $d(i, A_k) = \sum_{j=1}^{|\mathcal{R}|} d(i, j, A_k)$. The persistence statistic is then defined by $u(i, j, k, A_l, X) = d(i, j, A_l)/d(i, A_l)$, i.e., the fraction of i's outgoing communication volume which has been devoted to j. Where the associated θ parameter is positive, this effect produces a tendency for actors to preferentially direct action to those who have comprised the bulk of their past communication history. Such a phenomenon could emerge empirically from unobserved relational heterogeneity, as well as from cognitive processes such as the enhanced availability to memory of frequent communication partners. More broadly, a positive persistence parameter captures a tendency towards social "inertia," in a manner loosely analogous to the role played by a positive AR(1) term in an autoregressive time series process.

While it is most natural to think of persistence as a positive-sign effect, it is also possible to obtain negative persistence parameters. In this case, the model reflects a process of "partner switching," in which actors become less likely (ceteris paribus) to contact those who comprise a larger fraction of their past communication history. This could be induced by differential availability of actors over time, as well as by search processes (such as information seeking behavior) which encourage the accumulation of a diverse array of contacts.

2.2.3 Preferential Attachment

In the midst of a turbulent environment, judgments regarding potential communication targets may be highly uncertain. When one cannot be sure who is still able to respond, it is natural to utilize past communicative activity as a predictor of current availability: those who have been involved in past communication are more likely to be present and able to respond than those with no prior communicative activity. The phenomenon in which actors with a greater level of past activity are more likely to be chosen as communication targets is an example of *preferential attachment*, and is easily captured via a statistic of the form $u(i, j, k, A_l, X) = d(j, A_l) + \sum_{h=1}^{|\mathcal{S}|} d(h, j, A_l)$. Where the associated parameter is positive, actors with more past communication tend to become more attractive targets, creating a positive feedback loop which tends to lead to the creation of high degree actors. By contrast, a negative attachment parameter reflects a tendency to seek out actors who have not had prior involvement with the communication network, e.g., due to a novelty seeking process. In this case, the attachment process will tend to suppress the formation of high-degree actors, leading (ceteris paribus) to a "flatter" indegree distribution.

2.2.4 Recency

A substantial influence on communication generally (and radio communication in particular) is the practice of turn-taking. Because conversations generally involve multiple turns, a strong tendency exists for actors to direct communications towards those who have recently contacted them (thereby generating a sequence of call/response pairs). Such a process can be modeled within the present framework by a statistic such as $u(i,j,k,A_l,X) = \rho(i,j,A_l)^{-1}$, where $\rho(i,j,A_l)$ is j's recency rank among i's in-neighborhood. Thus, if j is the last person to have called i, then $\rho(i,j,A_l)^{-1} = 1$. This falls to 1/2 if j is the second most recent person to call i, 1/3 if j is the third most recent person, etc. (To ensure that the behavior of ρ is well-defined, actors who do not belong to i's in-neighborhood are considered to have rank ∞.)

Where the parameter associated with the recency statistic is positive, actors exhibit a tendency to preferentially call those who have most recently contacted them. By turns, a negative parameter value would indicate a tendency to avoid calling those with more recent incoming communications. Such an effect seems unlikely to emerge within a context such as radio communications, but might be observed for other types of relational events (e.g., dominance contests (Chase et al., 1998)).

2.2.5 Triadic Effects

The last category of effects considered here are those arising from triadic forms. In contrast with those properties already considered (which are, at best, dyadic), triadic effects engender dependencies which are far less local in nature (Frank and Strauss, 1986; Strauss, 1986). The most famous of these effects is that related to transitivity, which may be understood here as the tendency for the existence of one or more i,j two-paths to enhance or inhibit direct communication from i to j. The impact of the same two-paths on the corresponding j,i communications is naturally understood as a *cyclicity* effect, and may be motivated by the notion that the target of a brokered communication may be likely to bypass the broker when replying (thus forming a direct connection, and creating a cycle). Each of these effects may be parameterized via statistics of the form $u(i, j, k, A_l, X) = \sum_{h=1}^{|\mathcal{R}|} \min\{d(i, h, A_l), d(h, j, A_l\}$ and $u(i, j, k, A_l, X) = \sum_{h=1}^{|\mathcal{R}|} \min\{d(j, h, A_l), d(h, i, A_l\}$ (respectively). The associated parameters then simply indicate the strength of the tendency to form cycles/transitive closure, or to inhibit the same, depending on the sign of the parameter value.

In addition to the "classic" two-path effects, it is also useful to consider the potential impact of shared partners on direct interaction (Snijders et al., 2004). For instance, two actors who both have contacted the same third parties may be more or less likely to contact one another directly; this is referred to an *outbound shared partner* effect. Similarly, one can imagine an effect due to having been contacted by the same third party, which would constitute an *inbound shared partner* effect. These effects are, respectively, indexed by the statistics $u(i, j, k, A_l, X) = \sum_{h=1}^{|\mathcal{R}|} \min\{d(i, h, A_l), d(j, h, A_l\}$ and $u(i, j, k, A_l, X) = \sum_{h=1}^{|\mathcal{R}|} \min\{d(h, i, A_l), d(h, j, A_l\}$. As with the two-path effects, the sign and magnitude of the parameters associated with these statistics indicate the extent to which such configurations are encouraged or inhibited via the dynamic process.

2.3 Parameter Estimation

Given a choice of sufficient statistics, either Equation 2 (in the exact case) or Equation 3 (in the ordinal case) defines the likelihood of A_t under the relational event model. Since both expressions are readily computable, there is in principle no difficulty in carrying out likelihood-based inference for θ given A_t. (Indeed, the relative ease of likelihood computation for this model family dramatically reduces the difficulties frequently encountered with dynamic tie-based models such as those of Snijders (2005), or static ERGs (see, e.g. Handcock, 2003, for a description of some of these issues).) The most obvious tactic to employ in this regard is maximum likelihood estimation, i.e. identifying $\hat{\theta}$ such that

$$\hat{\theta} = \arg\max_{\theta} p(A_t|\theta) \tag{5}$$

using a variant of Newton-Rapheson, simulated annealing, or other heuristic optimization method (see Acton, 1990, for a number of approaches). Once $\hat{\theta}$ has been calculated, the inverse information matrix at the MLE can be employed to obtain approximate standard errors in the usual fashion. Alternately, fully Bayesian estimation of θ can be performed by positing a prior distribution on θ and maximizing and/or simulating draws from

$p(\theta|A_t) \propto p(A_t|\theta)p(\theta)$. Though we do not treat the issue in detail here, simulation of posterior draws for the relational event model is fairly straightforward using a Metropolis algorithm; see Gelman et al. (1995) for an overview of this appeoach.

One computational challenge which does emerge is the need to calculate the product of survival functions (or sum of rates, in the ordinal case) across all $|\mathcal{S}| \times |\mathcal{R}| \times |\mathcal{C}|$ possible events at each iteration. Where the dimensions of this product become large, the number of elements involved can quickly get out of hand; even given a single action type, the complexity of this calculation will generally grow with the square of the number of network members. In this case, it may be feasible to replace the relevant quantity with a Monte Carlo estimate, based on explicit calculation of a limited number of events. For instance, let $s'_1, \ldots, s'_m$, $r'_1, \ldots, r'_m$, and $c'_1, \ldots, c'_m$ be drawn uniformly from $1, \ldots, |\mathcal{S}|$, $1, \ldots, |\mathcal{R}|$, and $1, \ldots, |\mathcal{C}|$ (respectively). Then we may approximate the normalizing factor of Equation 3 by the Monte Carlo quadrature

$$\sum_{j=1}^{|\mathcal{S}|} \sum_{j=1}^{|\mathcal{R}|} \sum_{j=1}^{|\mathcal{C}|} \lambda_{jkl\tau_{i-1}} \approx \frac{|\mathcal{S}||\mathcal{R}||\mathcal{C}|}{m} \sum_{j=1}^{m} \lambda_{s'_j r'_j c'_j \tau_{i-1}} \tag{6}$$

(see, e.g. Kalos and Whitlock, 1986). Depending on the sufficient statistics on which λ depends, stratification of sender, receiver, and/or action type may be required to ensure convergence of the estimated likelihood. In particular, one should beware of any scheme which results in a failure to cover each sender, receiver, and action type (particularly where fixed effects are employed). Stratification can also reduce the variance of the estimator, which can be estimated by the appropriately scaled variance of the sampled rates. As a rule of thumb, the standard deviation of the estimated normalizing factor should be small compared to both $\lambda_{s_i r_i c_i \tau_{i-1}}$ and to the estimated normalizing factor itself. Where this condition is met for all i, the estimated likelihood will closely approximate the exact likelihood, and the resulting estimators should be well-behaved. Otherwise, it may be necessary to increase m and/or employ additional stratification so as to increase the precision of the estimated normalizing factor.

3 The WTC Data

To illustrate the use of the relational event framework, we here apply the ordinal data model to several documents from a larger collection released by the Port Authority of New York and New Jersey. Specifically, we analyze transcripts of radio communications among six groups of responders to the WTC disaster. Each transcript documents the voice communication associated with a single channel; each channel was used exclusively by a single group of responders.[3] The six transcripts employed here encode (in increasing order of length) the Port Authority Trans-Hudson channel 27 (PATH Radio), Newark airport maintenance (Newark Maint), Newark airport police (Newark Police), New Jersey State Police Emergency Network channel 2 (NJSPEN 2), Newark airport command post/dispatch (Newark CPD), and World Trade Center police (WTC Police) channels. Lengths range from 70 to 481 eligible transmissions (see below), with the number of named communicants ranging from 24 to 46. The time period covered by each transcript begins with the impact of the first plane into the North Tower at 8:46 AM, and ends at three hours and thirty-three minutes or (where relevant) until the collapse of the structures containing the communicants (roughly one hour and fifteen minutes).

3.1 Coding

Each radio transcript contains a list of transmissions exchanged among responders, presented in chronological order. Some sender information is provided by the transcriber; depending on the specific transcript, this includes some or all of name, rank, gender, and organization. This information, together with transcript content (including communicants' use of names and callsigns, sequence information, and conversatonal cues), was used to assign a unique identifier to the sender and named target(s) of each transmission. Where one-to-many communications were encountered, each was coded as a series of dyadic transmissions from the sender to each of the named recipients (in the order named). Transmissions with no clear target(s), and/or targets which were identified only as a group (e.g., "anyone," "all units") are outside the scope of person-to-person communications considered here, and were removed from the data set. The resulting lists of ordered transmissions (one per transcript) comprised the event sets (A_t) employed in subsequent analyses. The sets of potential senders and receivers for each transcript ($\mathcal{S}, \mathcal{R}$) were taken to be the union of named communicants from the transcript in question, with τ corresponding to the order of appearance for each

[3]Examination of transcript content, as well as other supporting materials (including the 9/11 Commission report (National Commission on Terrorist Attacks Upon the United States, 2004)) strongly suggests that the groups studied here lacked access to other radio channels. Thus, we treat the two as effectively equivalent for purposes of this study.

transmission event. Finally, because we analyze only radio communications, we consider all relational events to be of the same type (i.e., $|\mathcal{C}| = 1$).

In addition to the relational event data itself, we consider the formal status of individual responders as an illustrative covariate. Specifically, we attempted to identify individuals within each organization whose formal roles entailed coordinative responsibilities. As this was not directly available, such status was inferred from the content of the available transcripts. We coded communicants as occupying institutionalized coordinator roles if their transcriber-assigned labels or within-transcript terms of address contained one of the following words: "command," "desk," "operator," "dispatch(er)," "manager," "control," and "base." Within the Newark Airport transcripts, the content of the communications suggested that actors were referring to a centralized Newark Airport command desk as simply "Newark Airport," so this individual was also assigned to institutionalized status.

3.2 Model Selection

We begin our investigation of the WTC data by fitting a range of models using the effects enumerated in Section 2.2. In each case, parameter estimates were obtained using maximum likelihood under the ordinal time model; the latter was employed due to the fact that exact temporal information was not available for this data set. Size descriptives for the six transcripts treated here are shown in the first two lines of Table 1 . N refers here to the number of actors within the network, and M refers to the number of distinct communications recorded. (M is thus the most natural quantity measure for this data.)

Proceeding from the first two rows, all subsequent entries within Table 1 consist of BIC scores (Wasserman, 2000) arising from the indicated model/data combination. Models are listed by effects, with codes corresponding to fixed effects (FE), persistence (P), preferential attachment (PA), recency (R), and triads (T). The null model (listed in the eponymous row) treats all events as equiprobable, and thus serves as a reference for the other models. The next block of models (represented by single terms) include only one effect in each case, and can thus be intrepreted as providing evidence of marginal effects. Finally, the third block seeks to combine effects in a manner which facilitates the analysis of hub formation. For this purpose, recency and the triadic effects are taken as "controls," with the two major alternatives being the cumulative effects (P and PA) and the fixed effects (FE). By investigating BIC scores across models, we can thus evaluate the extent to which one mechanism versus another appears to be providing a

Network	PATH Radio	Newark Maint	Newark Police	NJSPEN 2	Newark CPD	WTC Police
N	28	25	24	26	46	35
M	70	77	83	149	271	481
Null	927.93	985.13	1048.05	1930.14	4138.34	6812.60
R	659.36	521.08	650.49	1431.95	2946.52	4081.38
T	941.95	999.79	1060.45	1780.55	4034.06	5853.89
P	755.99	702.57	786.26	1684.74	3796.24	5754.44
PA	902.86	901.04	1021.68	1711.58	3766.50	5703.66
FE	920.27	902.53	1041.14	1381.78	3337.86	4308.54
R+T	675.55	538.20	667.29	1329.46	2872.93	3685.39
R+T+P+PA	682.48	538.50	672.34	1326.76	2833.93	3639.47
R+T+FE	733.33	585.92	656.94	1197.55	2733.42	3486.46
P+R+T+PA+FE	740.61	581.60	654.44	1203.44	2710.91	3384.17

Table 1: Data Size and BIC Statistics for the Fitted Relational Event Models

more parsimonous account of the data.

Looking across the BIC values of Table 1, a number of patterns clearly emerge. The first, and most striking, is the strong impact of recency on structural dynamics: no preferred model omits the R term, and in some networks this effect alone generates the best fitting model. The opposite pattern is exhibited by the triad statistics, which had very little success in explaining most of the data sets. The cumulative (P and PA) terms clearly do have some impact on network dynamics, but that impact is fairly weak compared to recency effect; in neither case was either able to unseat the marginal R model as the more favored option. Interestingly, the fixed effect terms prove generally more powerful than the cumulative terms, particularly as the amount of data increases. This implies that unobserved heterogeneity is a more powerful influence within the data than cumulation of ties, suggesting the possibility that unmeasured aspects of the context of interaction may be strong determinants of relational dynamics. This is all the more striking because of the well-known conservatism of the BIC as model selection index, which tends to favor smaller models (Wasserman, 2000). Although the FE terms add a very large number of parameters to the model, they are still sufficiently successful at increasing the maximized likelihood to be considered viable under the BIC. Together with the cross-network consistency of fixed effect performance, this speaks strongly to the importance of latent heterogeneity in modeling networks like those used here.

3.3 Parameter Estimates

To get a better idea of the dynamics of communication at the WTC, we now turn to the estimated parameters for the relational event model. Figure 1 shows MLEs (and associated asymptotic 95% confidence intervals) for the marginal triadic (T-ISP, T-OSP, T-ITP, T-OTP), recency (R), persistence (P), and preferential attachment (PA) effects; although included in the same figure for comparison, each effect category was fit independently to each transcript.[4] Consistent with Table 1, we observe reasonably strong and systematic marginal effects for recency, persistence, and preferential attachment. All are positive, suggesting marginal tendencies towards reciprocity, persistence in selection of communication targets, and preferential targetting of actors with higher levels of prior communication activity. Triadic effects, however, are more varied: significant effects are observed for only

[4]95% confidence intervals are based on an asymptotic z approximation, with standard errors derived from the inverse Fisher information matrix at the MLE. Some confidence intervals have been truncated for clarity of display.

three out of the six transcripts, and little consistency is observed in strength or direction of effect. That said, the three shorter transcripts provide little information regarding triadic structure (as reflected in the large standard errors), leaving open the possibility of subtle effects beneath the detection threshold of the model. While we thus cannot rule out the possibility of consistent triadic biases in the WTC communication networks, the data does not provide evidence in support of this assertion.

While the marginal effects shown in Figure 1 appear intuitive, they may also be misleading: real social systems involve multiple, interacting mechanisms, whose joint effects can be nonobvious. As Table 1 indicates, the BIC-preferred models for all but the shortest transcripts incorporate multiple effects, which must be considered jointly in order to obtain realistic estimates. With this in mind, Figure 2 once again shows parameter estimates for triadic, recency, persistence, and preferential attachment effects, this time from a joint model in which all effects are included (fixed effects entered, but not shown). These estimates indeed paint a very different picture: once heterogeneity in activity level is controlled for, the "cumulative" mechanisms of persistence and preferential attachment either lose significance or (with one exception) reverse direction. This strongly suggests that the primary impact (if any) of cumulative mechanisms in the WTC communication system is to induce *partner switching*, rather than concentration on a small number of hub nodes. It follows, then, that hub formation must in general be a consequence of latent heterogeneity at the actor level. Although this result illustrates the subtleties which can be rendered visible by the joint consideration of multiple mechanisms, it should also be noted that some parameters do not change substantially in comparison with their marginal estimates. A strong, positive recency effect continues to be observed for all networks, underscoring the robust importance of dynamic reciprocity within the WTC communication structure. By turns, triadic biases continue to be minimal, with most already weak effects becoming insignificant once other mechanims are controlled for. The New Jersey State Police Emergency Network does show some remaining tendency to discourage communications between responders with shared references and/or which create cycles, but these effects are fairly small in practice. While some triadic biases do exist, then, they do not seem to be a major driving force within the communication system.

Clearly, the above analyses suggest that responder-level heterogeneity in activity levels (as captured by the fixed effect parameters) plays an important role in determining communication network structure. While this heterogeneity could stem from many sources – including differences in con-

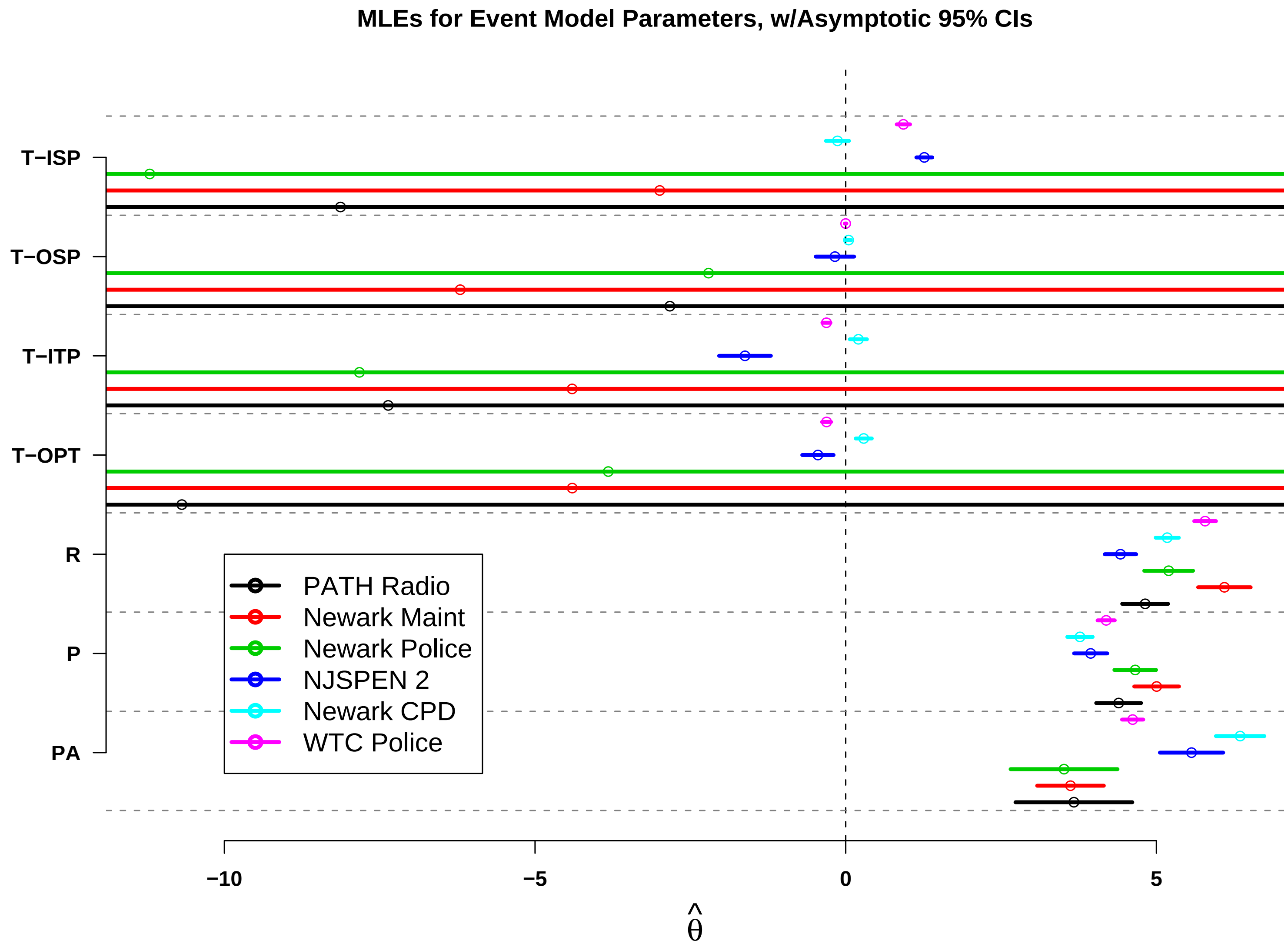

Figure 1: Parameter Estimates and Approximate 95% Confidence Intervals, Marginal Models

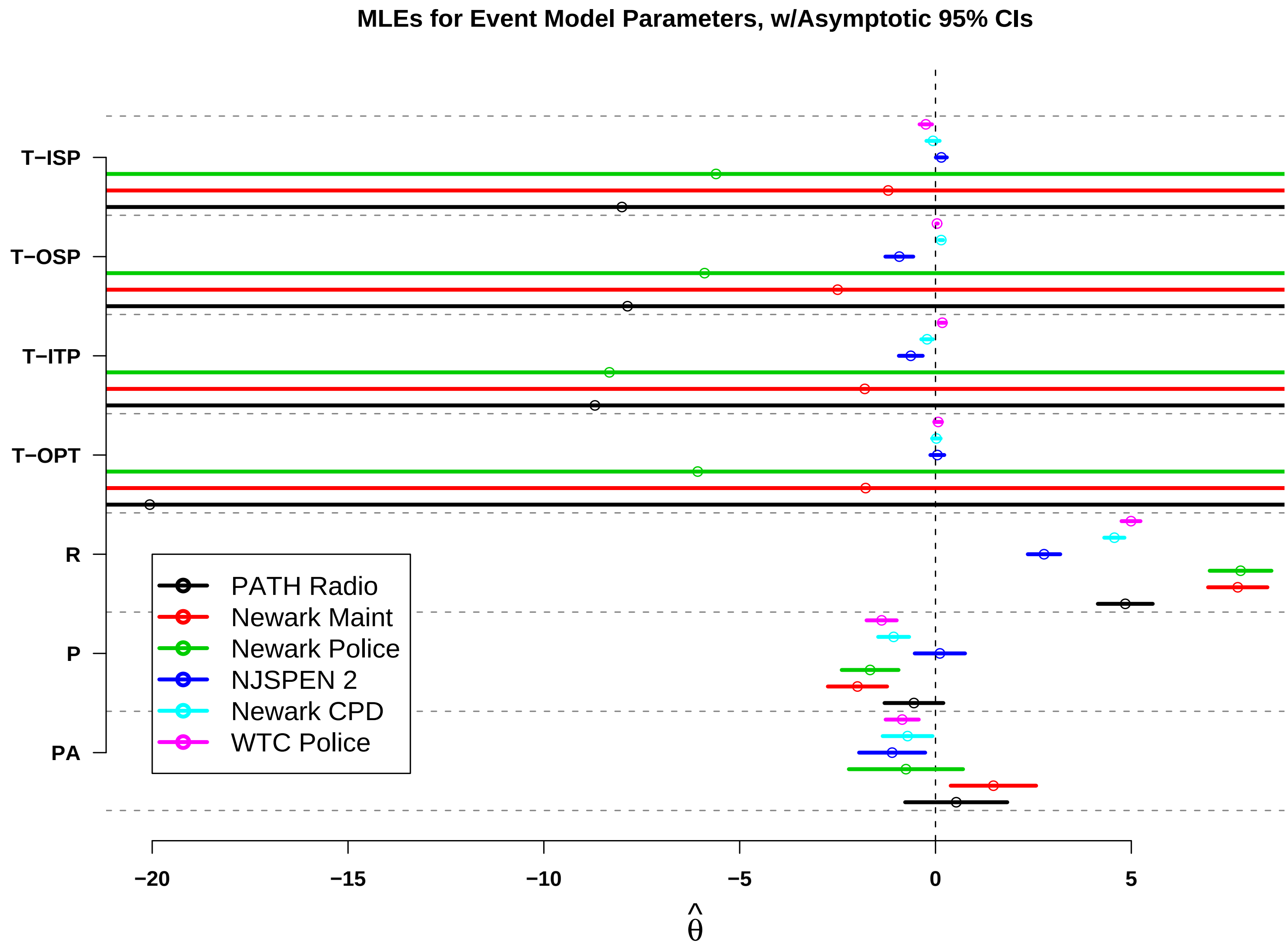

Figure 2: Parameter Estimates and Approximate 95% Confidence Intervals, Joint Models

text, training, or cognitive/emotional state – we will here consider only the possible influence of institutionalized coordinative roles. Intuitively, one may expect responders with such roles to act as hubs within the communication network, thus displaying higher levels of activity (net of other processes) than actors without such roles. On the other hand, demands for emergent coordination (Dynes, 2003) may render such roles largely irrelevant during the immediate aftermath of a high-consequence event. To assess this possibility, we compare fixed effect estimates under the joint model for responders holding institutionalized coordinative roles with those for other responders. Figure 3 shows boxplots for activity level effects by institutional status, for all six WTC networks. As the figure suggests, the impact of institutionalized coordinative roles is weak at best: medians for the institutionalized coordinators are higher in only four out of the six networks, and the pooled mean difference is not significant once parameter uncertainty is taken into account ($z = -1.42$, $p = 0.16$). Our analysis thus reinforces the findings of Petrescu-Prahova and Butts (2005), whose static analysis of WTC communications found evidence that centrality within responder radio communication networks was largely due to factors other than institutional status.

4 Summary and Conclusion

In this paper, we have introduced a stochastic model for intertemporal activity data, based on a relational event formalism. Using the samplifying assumption of piecewise constant hazards, we are able to construct a fairly broad modeling framework which can be applied to data with either exact or ordinal timing information. A wide range of mechanisms can be evaluated within this framework, of which several are illustrated here; once specified, parameters associated with the direction and strength of these effects can be readily estimated using maximum likelihood.

To summarize our substantive findings, our analysis of six transcripts from the World Trade Center disaster suggests that heterogeneity in base activity level – not persistence or preferential attachment – is the key driver in the formation of hubs within WTC radio communication networks. As expected, strong reciprocity effects were observed for all six transcripts; triadic effects, on the other hand, did not appear to play a large role in driving communication dynamics. Although we cannot currently identify the source of the responder heterogeneity observed here, our analysis indicates that institutionalized coordinative roles have little explanatory power in this regard. This is consistent with the hypothesis that emergency phase responder ac-

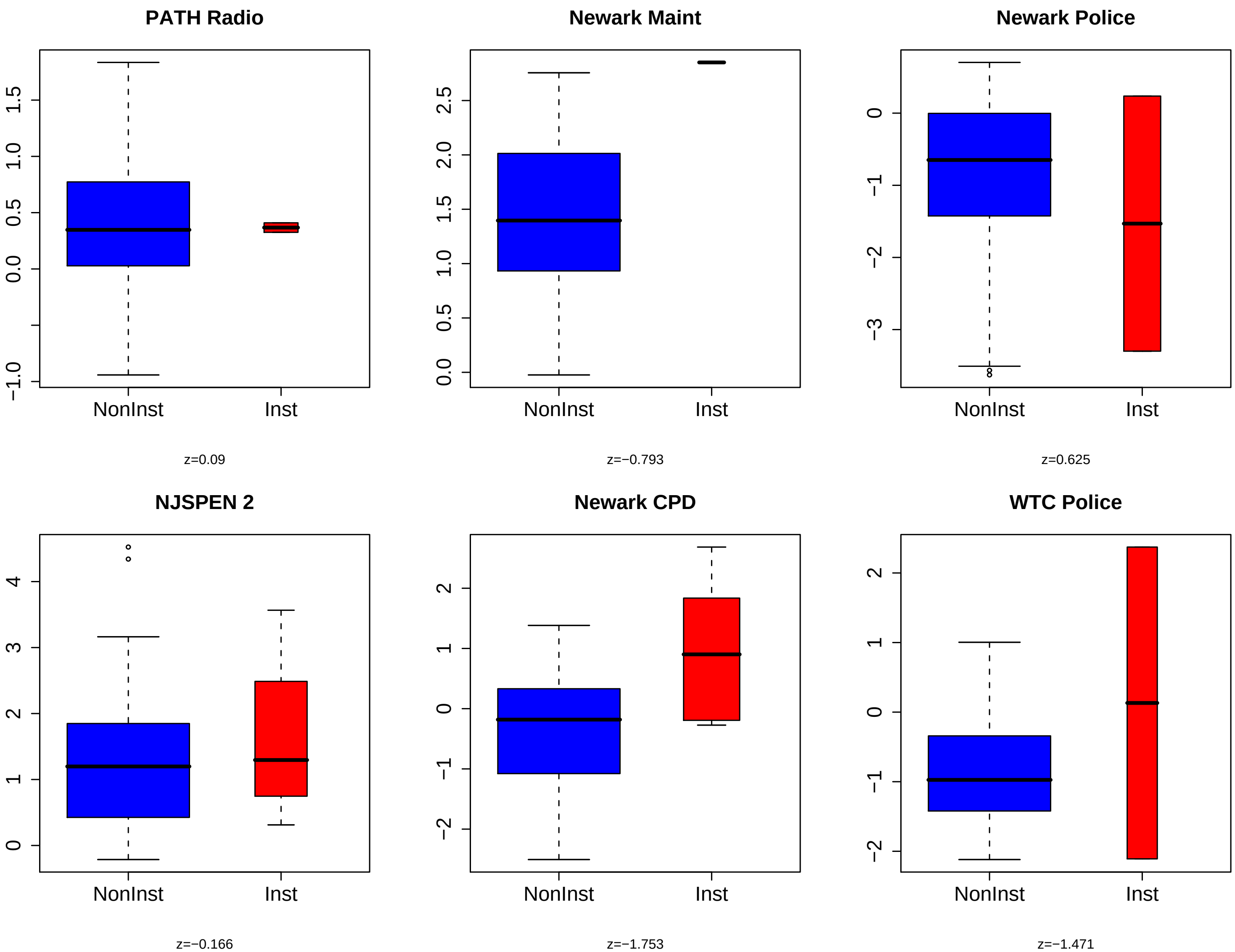

Figure 3: Fixed Effect Parameters by Institutional Status, Joint Models

tivity at the WTC was largely driven by idiosyncratic, situational factors which overwhelmed prior organization. This is broadly consistent with the findings of the 9/11 Commission (National Commission on Terrorist Attacks Upon the United States, 2004), though it should be cautioned that alternative explanations may also exist. Additional analyses with a larger body of data should shed further light on this issue.

5 References

Abe, K. (1976). The Behavior of Survivors and Victims in a Japanese Nightclub Fire: A Descriptive Research Note. *Mass Emergencies*, 1:119–124.

Acton, F. S. (1990). *Numerical Methods That Work*. Mathematical Association of America, Washington, DC, third edition.

Blossfeld, H. P. and Rohwer, G. (1995). *Techniques of Event History Modeling: New Approaches to Causal Analysis*. Lawrence Erlbaum and Associates, Mahwah, NJ.

Chase, I. D., Tovey, C., Spangler, D., and Manfredonia, M. (1998). Individual attributes do not explain linearity in dominance hierarchies: A behavioral and evolutionary puzzle. Working paper, SUNY Buffalo.

Coleman, J. S. (1990). *Foundations of Social Theory*. Harvard University Press, Cambridge, MA.

Drabek, T. E. (1986). *Human System Responses to Disaster: An Inventory of Sociological Findings*. Springer-Verlag, New York.

Dynes, R. R. (1970). *Organized Behavior in Disaster*. Heath Lexington, Lexington, MA.

Dynes, R. R. (2003). Finding Order in Disorder: Continuities in the 9-11 Response. *International Journal of Mass Emergencies and Disasters*, 21(3):9–23.

Frank, O. and Strauss, D. (1986). Markov graphs. *Journal of the American Statistical Association*, 81:832–842.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (1995). *Bayesian Data Analysis*. Chapman and Hall, London.

Goffman, E. (1963). *Behavior in Public Places: Notes on the Social Organization of Gatherings*. Free Press, New York.

Handcock, M. S. (2003). Assessing degeneracy in statistical models of social networks. CSSS working paper no 39.

Holland, P. W. and Leinhardt, S. (1981). An Exponential Family of Probability Distributions for Directed Graphs (with Discussion). *Journal of the American Statistical Association*, 76(373):33–50.

Kalos, M. H. and Whitlock, P. A. (1986). *Monte Carlo Methods, Volume I: Basics*. John Wiley and Sons, New York.

Macy, M. (1991). Chains of cooperation: Threshold effects in collective action. *American Sociological Review*, 56:730–747.

Mileti, D. S., Drabek, T. E., and Haas, J. E. (1975). *Human Systems in Extreme Environments*. Institute of Behavioral Science, University of Colorado, Boulder, CO.

Miller, J. H., Butts, C. T., and Rode, D. C. (2001). Communication and cooperation. *Journal of Economic Behavior and Organization*, 46(4).

National Commission on Terrorist Attacks Upon the United States (2004). The 9/11 Commission report: Final report of the National Commission on Terrorist Attacks Upon the United States. New York.

Noji, E. K. (1997). Earthquakes. In Noji, E. K., editor, *The Public Health Consequences of Disasters*, pages 135–178. Oxford University Press.

Petrescu-Prahova, M. and Butts, C. T. (2005). Emergent coordination in the World Trade Center disaster. IMBS Technical Report MBS 05-03, Institute for Mathematical Behavioral Sciences, University of California, Irvine, Irvine, CA.

Quarantelli, E. L. (1960). Images of Withdrawal Behavior in Disasters: Some Basic Misconceptions. *Social Problems*, 8:68–79.

Snijders, T. A. B. (2005). Models for longitudinal network data. In Carrington, P. J., Scott, J., and Wasserman, S., editors, *Models and Methods in Social Network Analysis*, pages 215–247. Cambridge University Press, New York.

Snijders, T. A. B., Pattison, P. E., Robins, G. L., and Handcock, M. S. (2004). New specifications for exponential random graph models. CSSS working paper no 42.

Strauss, D. (1986). On a General Class of Models for Interaction. *SIAM Review*, 28(4):513–527.

Wasserman, L. (2000). Bayesian model selection and model averaging. *Journal of Mathematical Psychology*, 44(1):92–107.