Most diffusion models make the implausible assumption that diffusion is unstructured by relationships within the population of interest. This article proposes methods for incorporating a priori notions about social structure into analysis. Diffusion is modeled within an event history framework where the individual's rate of adoption is a function of prior adoptions by related actors. Two diffusion models are suggested: an epidemic model where adoption rates vary with the number of prior adoptions, and a salience model where adoption rates vary with time since the last event. This approach is illustrated in an examination of the decolonization of British and French colonies. Diffusion is shown to occur within regions rather than within empires or the world system as a whole.

Adding Social Structure to Diffusion Models

An Event History Framework

DAVID STRANG
University of Iowa

Mathematical models of diffusion commonly assume a homogeneous mixing population, where transmission between each pair of prior and potential adopters is equally likely. But sociologists examining diffusion are often interested precisely in the way this assumption does not hold—in how relations between members of a population channel diffusion. Such an inquiry does more than make an analysis of diffusion more realistic. It turns it into a search for social structure.

This article presents methods for incorporating the social structure of a population into diffusion analyses. Event history methods provide a framework within which individual-level analogues of classical diffusion models may be formulated. Models posed at the individual level allow the analyst to investigate complexities in the relation between social structure and diffusion that are intractable at the population level.

The first section considers the relation between social structure and diffusion. The second proposes two models of diffusion and reviews estimation procedures. The third illustrates these models for the diffusion of sovereignty among colonial dependencies in the twentieth century.

AUTHOR'S NOTE: *I would like to thank Carol Conell, Scott R. Eliason, John W. Meyer, Nancy B. Tuma, and the editor and reviewers of Sociological Methods & Research for their comments on earlier versions of this article.*

SOCIOLOGICAL METHODS & RESEARCH, Vol. 19, No. 3, February 1991 324-353

SOCIAL STRUCTURE AND DIFFUSION

The term *diffusion* is used here to refer to any process where prior adoption of a trait or practice in a population alters the probability of adoption for remaining non-adopters. This definition is more general than one referring to direct contact between prior adopters and potential adopters (though I use the metaphor of actor-to-actor transmission in describing these processes). The notion of diffusion is central to the study of epidemics and the spread of information or rumors. It has also been applied to sociological concerns such as the spread of drug prescription practices among physicians (Coleman, Katz, and Menzel 1966), race riots in urban areas (Spilerman 1970), and civil disorders in nation-states (Pitcher, Hamblin, and Miller 1978). In this article I will refer to the adoption acts of "actors" for simplicity, where actors may be individuals, groups, organizations, or national polities.

Though mathematical treatments of diffusion generally assume homogeneous mixing (Bailey 1975), sociological research has often sought to "bring social structure back in." Coleman (1964) discussed several approaches to analyzing diffusion in "incomplete social structures" where homogeneous mixing does not hold. Spilerman (1970) entertained the hypothesis of geographic contagion in race riots in the 1960s. Burt (1987) investigated the impact of linkages based on cohesion and structural equivalence in a reanalysis of Coleman, Katz, and Menzel's classic *Medical Innovation*. For a more general discussion of spatial correlation, see Doreian (1981).

Despite differences in analytic procedures, the use of "social structure" in these studies is the same. In each, the analyst defines social relations which link actors to each other. Such relations comprise a

social structure that "channels" diffusion, so actors respond to those they are socially connected to. The article also adds social structure to diffusion in just this fashion.[1]

Methodological tools for the analysis of heterogeneity in diffusion are not well developed, however. This can be seen in the limitations of each of the studies mentioned above. Coleman's approach required that social structure take the form of a partition of the population. Spilerman asked whether the geographic clustering of racial disturbances could be explained by individual city characteristics, an indirect test of diffusion at best. Burt regressed the date physicians adopted tetracycline on weighted means of the adoption dates of other physicians with network proximities as the weights. This procedure permits future events to predict present ones, violating a basic requirement of causal analysis. Further, regression-based approaches do not deal naturally with those members of a population who are right censored (fail to adopt before the end of the observation period).

The event history methods presented below overcome these limitations. They are able to handle social structures measured in very different ways, from partitions of the population to proximity matrices. And they are ideally suited for the analysis of a temporal process, ensuring appropriate temporal sequences and handling of censored observations. Further, an event history framework permits ways of modeling the interdependence among events that go well beyond classic diffusion models. This article makes a first step in this direction, presenting some extensions of standard event history methods that capture simple effects of social structure.

A particular strategy to modeling the connections between diffusion and social structure is emphasized in this article. I suggest that a useful approach is to contrast the impact of several social relations. There are often multiple social relations that may be hypothesized to channel diffusion. By counterposing these effects, the analysis becomes something more than a specification of the path of diffusion; it arbitrates between competing notions of social structure.

Burt's (1987) analysis exemplifies the promises of this strategy (though again I have concerns about the methods employed). He contrasted the extent to which diffusion worked along the lines of

structural equivalence (doctors were influenced by those who had similar relations to other doctors) versus the lines of social cohesion (doctors were influenced by those they were directly connected to). The finding that structural equivalence seemed to channel diffusion, while cohesion did not, speaks to an important debate within social network theory.

MODELS

Diffusion processes are generally modeled in terms of the rate of change in the size of the adopting population. But the incorporation of even simple patterns of linkages between individual actors quickly makes a population-level analysis intractable. I adopt an individual-level analysis instead. Individual adoption rates are specified as functions of the prior adoption acts of related actors. Event history methods (Tuma and Hannan 1984) provide a natural framework within which appropriate models may be stated and estimated.

AN EPIDEMIC MODEL OF DIFFUSION

I will call the standard formulation of diffusion the epidemic model, where adoption depends on transmission between prior and potential adopters, and the contagiousness of prior adopters (transmitters) is assumed invariant over time. A *population level* stochastic formulation of the epidemic model, assuming homogeneous mixing, is

$$\lim_{\Delta t \downarrow 0} \frac{Pr(N(t+\Delta t) = n+1 \mid N(t) = n)}{\Delta t} = [\alpha + \beta n(t)]\,[P - n(t)] \qquad (1)$$

$N(t)$ is a random variable giving the number of adopters at time t and $n(t)$ is its realization. P is the size of the population. Thus $Pr(\cdot)$ is the probability that a population of size P moves from n adopters at time t to $n+1$ adopters at time $t+\Delta t$.

To indicate the logic underlying (1), the right-hand side may be decomposed into the contributions of diffusion from outside and inside the population. The impact of diffusion from an external source is

proportional to the number of possible contacts which under homogeneous mixing is the number of prior adopters $n(t)$ multiplied by the number of individuals at risk $P - n(t)$. With α and β the respective intensities, the contribution of diffusion from outside the population is thus $\alpha[P - n(t)]$, and the contribution of diffusion from inside is $\beta n(t)[P - n(t)]$.

This kind of nonlinear stochastic model is difficult to work with. In fact, a general solution has not been discovered for (1). For large P when $n(t)$ is far from equilibrium, Bartholomew (1982) demonstrates that

$$E[N(t)] = \frac{(e^X - \alpha/\beta + 1)P}{e^X + P} \tag{2}$$

where $X = \beta(P - \alpha/\beta)t - \gamma + \phi(\alpha/\beta - 1)$.[2] This sigmoid curve lies above a logistic curve, which is the solution for a deterministic version of the model. The two are only equal under quite restrictive conditions (see Bartholomew 1982, pp. 255-59).

My central contention is that progress in developing models of heterogeneous mixing can be made by recasting the diffusion process at the individual level. This strategy simplifies the model and allows variations in social structure to be introduced. To this end, population dynamics may be restated at the *individual level* in terms of *instantaneous transition rates* or *hazards*

$$r_i(t) = \lim_{\Delta t \downarrow 0} \frac{Pr(Y(t + \Delta t) = y_1 \mid Y(t) = y_0)}{\Delta t} \tag{3}$$

$r_i(t)$ gives the limiting probability that an individual case moves from non-adoption (state y_0) to adoption (y_1) between t and $t + \Delta t$. Transition rates are indexed by i throughout to emphasize that they describe individual movement.

To translate the right-hand side of (1) into a corresponding model for (3), note that the population level equation may be thought of as multiplying the probability of each individual's adoption, which is $\alpha + \beta n(t)$, by the number of individuals who are at risk of adopting, $P - n(t)$. When we move to an individual level formulation, we have $P - n(t)$ transition rates of the form

$$r_i(t) = \alpha + \beta n(t) \tag{4}$$

This equation embodies the assumption of homogeneous mixing; rates are a function of the number of adopters in the population at time t, $n(t)$, since transmission can occur between any prior adopters and each potential adopter.[3]

It is now possible to introduce heterogeneous mixing. To substitute the assumption that diffusion flows along the lines of one or more social relations, equation (4) may be re-expressed as

$$r_i(t) = \alpha + \sum_{k=1}^{K} \beta_k n_{ik}(t) \tag{5}$$

where k indexes and K different social relations. $n_{ik}(t)$ gives the number of prior adopters at time t in the kth relation for case i. This is now the relevant number of prior adopters who can transmit to a potential adopter. For example, if gender is the sole relation structuring diffusion, then adoption rates for men are a function of the number of prior male adopters, while adoption rates for women are a function of the number of prior female adopters. If friendship is a second relation structuring diffusion, an individual's adoption rate is also a function of the number of his or her friends who have already adopted.

The article focuses on models that can be expressed by (5), where multiple social relations are hypothesized to channel diffusion, and each social relation identifies a set of potential transmitters for each adopter. Since the nonnegativity of transition rates is conventionally assured by expressing the rate as an exponential function, estimated models take the form

$$r_i(t) = \exp[\alpha + \sum_{k=1}^{K} \beta_k n_{ik}(t)] \tag{6}$$

Note that the standard interpretation of α in a diffusion framework (diffusion from an external source) may be broadened to include all

factors that affect the rate of adoption in the same way across cases and time (the usual interpretation of α in event history analysis).

While not pursued in the illustration below, several extensions of equation (5) may be suggested. First, exogenous characteristics of individual cases may be added to the model

$$r_i(t) = \exp[\alpha_i + \sum_{k=1}^{K} \beta_k n_{ik}(t)] \quad (7)$$

In other words, α_i may be a parameter vector multiplying a vector of measured covariates, rather than simply a constant.[4] This extension is not emphasized here since it does not pose methodological issues beyond those involved in examining diffusion in isolation. But the ability to combine the analysis of diffusion with other kinds of factors — again impractical at the population level — is an important additional advantage of an individual level approach.

Equation (5) may be generalized in another direction to handle arbitrary structures of proximity within a population. To do so it is helpful to shift notation and decompose $n_{ik}(t)$. Let x_{ijk} index the proximity of two individuals i and j according to relation k. Define $y_j(t)$ to equal 1 if case j has adopted by time t, and 0 otherwise. Then

$$dr_i(t) = \alpha + \sum_{k=1}^{K} \sum_{j=1}^{P} \beta_k x_{ijk} y_j(t) \quad (8)$$

Again, the right-hand side may be exponentiated to ensure nonnegativity. Note that if x_{ijk} only takes on the values 0 and 1, x_{ijk} y_{jk} may be replaced by n_{ik}.

A SALIENCE MODEL OF DIFFUSION

While an event history framework handles the standard assumptions of the epidemic model, it also suggests another approach. Much work on diffusion models has explored variations in contagiousness; see Bartholomew (1982, pp. 272-320) for a review. Here, I propose a "salience" model as a related extension. While the guiding metaphor

for epidemic models is direct transmission between adopters and non-adopters, a salience model is motivated by a view of the potential adopter as a decision maker taking the acts of others into account.

Two simplifying assumptions about the monitoring process are useful. First, the potential adopter is only affected by the last adoption event. Second, the salience of the last event decays over time. Some research indicates that these assumptions may be useful. Zielske and Henry (1980) found that the memory of a stimulus declines exponentially since the last exposure to it, regardless of the number of prior exposures. Conell and Cohn (1989) found that the rate of French coal mining strikes declines exponentially since the time of the last strike within the *department.*

The notion of salience may be modeled for a homogeneous mixing population as

$$r_i(t) = \exp(\alpha) + \exp(\beta + \gamma t_e) \quad (9)$$

where t_e is time since the most recent event in the population. This is referred to in the literature as a Makeham model. At the time of the event the rate is $\exp(\alpha) + \exp(\beta)$; it declines to $\exp(\alpha)$ at a rate governed by γ (assuming γ is negative).

For diffusion channeled by social relations,

$$r_i(t) = \exp(\alpha) + \exp(\beta + \sum_{k=1}^{K} \gamma_k t_{ik}) \quad (10)$$

where t_{ik} is time since the last adoption in the kth relation of case i. As in the epidemic model, it is straightforward to add individual characteristics as exogenous covariates (α becomes α_i) and to build in dependence on historical time or age.

In this model, social linkages govern the values of the "clock" and the speed of decay from the moment an event occurs. The rate at the time of the event $[\exp(\alpha) + \exp(\beta)]$ and the asymptoic value approached as time since the last event grows large $[\exp(\alpha)]$ are not indexed by social linkages. For this reason, it is potentially misleading to analyze multiple social relations simultaneously. A better strategy

may be to first examine the effects of each social relation in separate equations and compare values of α, β, and γ across equations. As a second step one may then incorporate those social relations with comparable α's and β's in a single model and estimate several γ_k's.

It should also be noted that salience models cannot handle the variety of social structures manageable within epidemic models. One reason is that for t_{ik} to be defined, the first event within relation k for case i must have already occurred. In other words, the analyst must discard those periods "at risk" that are prior to first events.[5] For practical purposes, this is not much of a problem when the postulated social structure partitions a population into a few large, cohesive groups. But it makes it difficult to study a sparse network structure (such as friendship relations) and impossible to analyze social relations expressed as continuously varying proximities.

Despite these drawbacks, a salience approach is useful because it affords a distinct perspective on how diffusion might operate. While the epidemic model implies that prior events have a cumulative impact that does not decay over time, the salience model assumes that prior events have a noncumulative effect that decays with time. The empirical finding that one model fits but the other does not suggests qualitative differences in the nature of the diffusion process.

Epidemic and salience models provide only a first step in analyzing interdependencies within an event history framework. It may be useful to suggest what an additional step might look like. Epidemic and salience models are in a sense complementary; the first emphasizes the cumulation of prior events, while the second emphasizes variation over time in the impact of events. It would seem desirable to combine these insights and allow each prior event to have an impact that declines over time. Such an extension is, however, well beyond the scope of this article.

ESTIMATION

One reason for limiting this article to epidemic and salience models is that these may be estimated by standard maximum likelihood routines for event history data. They require only the division of

observation times into appropriate intervals and the creation of variables, as described below. By contrast, more complex formulations (such as the combined epidemic-salience model suggested above) would require new estimation routines.

For both epidemic and salience formulations, the analyst would begin by defining possible diffusion channels. Let us assume that each social relation specifies some set of members of the population as the relevant transmitters for each potential adopter. In the most complex case, each individual would have a unique set of possible transmitters. In the analysis performed below, the population is partitioned into cohesive groups, where potential adopters in the partition share the same transmitters. While data management shortcuts exist for the second case that are not available for the first, the two are not fundamentally different.

To explain the procedures involved, it may be helpful to first describe the data structure of event history analysis employed here. The basic unit is the spell, an interval of time during which an individual case is observed. Spells are defined such that an event either occurs at the end point of the spell or does not occur at all during the spell (in which case the spell is said to be censored). Each individual case history may be divided into as many spells as are convenient. In general, spells are defined to match known variation in explanatory variables, so each spell may be characterized by a single value on each explanatory variable.

For both epidemic and salience models, the trick is to construct spells that do not overlap the adoption times of potential transmitters for that case. That is, spells for each individual are subdivided according to the adoption times of those population members linked to the individual by social relations of interest. This allows each spell to be characterized by appropriate values of n_{ik} or t_{ik}. Given the construction of these variables, the diffusion models described above may be estimated by standard Gompertz (for epidemic models) or Makeham (for salience models) routines.

Appendix A illustrates the data manipulation involved, for a hypothetical dataset of six cases (dataset 1). It employs coding conventions appropriate for RATE, which should translate fairly directly into

conventions used in other event history software. The variables characterizing each spell are: the time at the beginning of the spell (ST), the time at the end of the spell (ET), the state occupied at the beginning of the spell (SS), and the state occupied at the end of the spell (ES). An event occurs if SS and ET differ; if they take the same values the spell is censored. There are two other variables in the data set: a case identifier (ID) and a variable partitioning the cases into two groups (G).

Dataset 2 contains the same data, but restructured to permit a diffusion analysis. This restructuring is based on the variable G as the hypothetical "diffusion channel" which defines two groups (A and B) within which diffusion is thought to operate. The original six spells have been divided into 12. For example, case 4 is now represented by three spells: one from 1918 to 1934, which is the time of the first event in its group (A); the second from 1934 to 1945, when the second event within group A occurs; and the third from 1945 until 1973, when case 4 itself has an event. Note that case 3, which is censored in 1951 (and thus never has an event), does not contribute to the redivision of spells. Also note that cases 5 and 6, which are in group B, are not affected by the timing of events in group A, and vice versa.

In addition to the redivision of spells, dataset 2 also includes three variables to record information about prior events. $n(t)$ counts the number of events within the case's group that occurred prior to the spell. Since spells have been subdivided to not overlap relevant prior events, $n(t)$ is also the number of events that have occurred within the case's group *at any time during the spell.* For example, no events have occurred within group A during case 2's first spell, from 1918 to 1934; from 1934 to 1945, one event has occurred within case 2's group.

The variables ST_e and ET_e in dataset 2 give the starting and ending times of each spell, measured since the previous event within the group. Again, this data is sufficient to characterized the spell precisely because by construction no events occur within groups during spells. Also note that these variables are undefined during the first spells of each case, as discussed above. In an analysis where time is measured since the last event, each of these spells would be discarded, effectively starting the analysis in 1934 for cases in group 1, and in 1952 for cases in group 2.

The procedure for handling additional diffusion channels is identical to the one described above. One can apply the spell division and variable creation process to a dataset repeatedly, without regard to the number of times the procedure has already been applied.

However, when more than one social relation is studied, the estimation of equation (10) involves some additional complications. As Tuma and Hannan (1984, pp. 196-97) note, all measures of time may be expressed in the same metric. If we have several measures of time t_{ik}, each may be written in terms of a single measure arbitrarily chosen from the set (call it t_{i1}) as follows:

$$t_{ik} = t_{i1} + (t^*_{i1} - t^*_{ik}) \tag{11}$$

where the t^*'s represent the dates of the two events. For example, imagine that for a doctor in Burt's (1987) study the most recent adoption by a structurally equivalent actor occurred in month 12, while the most recent adoption by a connected doctor occurred in month 15. Then until the next events within either of these groups, time since adoption by a connected actor equals time since adoption by a structurally equivalent actor minus three (12 – 15).

Hence we can rewrite (10) as

$$r_i(t) = \exp(\alpha) + \exp(\beta + \gamma_0 t_{i1} + \sum_{k=1, j=1}^{K} \gamma_k(t^*_{i1} - t^*_{ik})] \tag{12}$$

where $\gamma_0 = \sum_{k=1}^{K} \gamma_k$. All coefficients may be recovered by estimating (12) as a Makeham model, where t_{i1} is the clock and all other variables are fixed during spells. γ_1 may be recovered by subtracting the values of the estimated γ_k's from γ_0. To recover standard errors, the equation can be rewritten in a second time scale so γ_1 may be estimated explicitly.

A NOTE ON UNOBSERVED HETEROGENEITY

It is well known that unobserved heterogeneity produces observed time dependence even when rates are in fact constant over time

(Blumen, Kogan, and McCarthy 1955). Cases with higher transition rates tend to have events early, so the population at risk is increasingly made up of cases with low transition rates. Unobserved heterogeneity thus characteristically generates an observed decline in transition rates over time. Such a process should also produce a clustering of events for cases with similar attributes over time, as would the structured diffusion processes discussed in this article. How then can unobserved heterogeneity and diffusion be distinguished?

A fundamental clue lies in the pattern of time dependence—the way the rate varies over time. An epidemic model implies that rates *increase* monotonically with time. Equation (2) gives the expected number of adopters over time under homogeneous mixing; while not identical to the logistic curve, it has the same qualitative features. The population level rate is low early in the process, when there are few transmitters, and low late in the process, when there are few potential adopters. The corresponding individual-level hazard is monotonically increasing with time, since the number of transmitters $n(t)$ is monotonically increasing. Non-adopters in a population full of transmitters are individually quick to adopt, though the total number of adopters per unit time will be low.

By contrast, a salience model implies that rates are globally *constant* over time. Each new event resets the clock, so the expected inter-arrival time (time between events) is constant. While rates are constantly varying with time, no long-term trend is expected.

One may thus judge the initial plausibility of each approach by examining how rates vary with time. If they decline, then clustering probably reflects heterogeneity within the population. If they increase, then an epidemic model is suggested. If there is substantial clustering but little pattern over time, a salience model seems plausible.

Of course, this is only a first step. True time dependence could take any form, producing any possible observed pattern of variation in transition rates over time. For example, a combination of unobserved heterogeneity and a historically increasing transition rate might resemble an epidemic model. To examine this possibility, one may add exogenous variables to explicitly control for heterogeneity in the population.

DECOLONIZATION AS A DIFFUSION PROCESS

The above models are illustrated in an analysis of British and French decolonization between 1918 and 1987. In the eyes of many observers, prior decolonization produced widely held beliefs that decolonization was not only possible but inevitable, and the maintenance of dependencies by force illegitimate. Newly sovereign states encouraged the decolonization of remaining dependencies through the force of their example, material support, and their contribution to world opinion in settings like the United Nations (Nogueira 1963; Holland 1985).

Though introduced to illustrate the uses of the above framework, what follows is also a serious attempt to model the decolonization process. It falls short of a full analysis, since only the diffusion aspects of decolonization are treated. I would not wish to argue that diffusion is the sole factor in decolonization or that estimated diffusion effects might not differ with the inclusion of important exogenous variables into the model. But I would contend that diffusion plays an important role in twentieth century decolonization and that the following analysis may yield some insight into the factors involved.

DATA

The population under study is all French and British dependencies in existence between 1918 and 1987. Independence as a newly sovereign state and full incorporation into a sovereign state are treated as equivalent "adoptions" of sovereignty (I will avoid the use of the word *independence* to prevent confusion between its political and probabilistic meanings). The acquisition of sovereignty is coded to the day and month of the formal change in status. This precision is important in an analysis where counts or times of previous events serve as explanatory variables.[6]

I chose 1918 as a starting date since it marks a major shift in the international states system occurring shortly before the beginning of the twentieth-century wave of decolonization.[7] The goal was thus to control for unobserved heterogeneity in environmental conditions so as to focus on patterns of diffusion. In general, the major methodolog-

ical criterion for the starting date in a diffusion analysis is that it predate the first event (again, to estimate a salience model one must discard observations prior to relevant first events). My experience with the epidemic model indicates parameter estimates are quite robust with respect to the choice of the starting date, assuming it to be equal to or less than the date of the first event. I chose 1987 as an ending date to capture as much of the historical process as possible.

Ninety-six British and French colonies moved from formal dependency to sovereignty during this period. The first was Weihaiwei, a British enclave returned to China in 1930. The height of decolonization was attained in 1960, when three British and thirteen French colonies became new nations. The most recent event occurred in 1983, when the Caribbean islands of St. Kitts and Nevis became a sovereign state. A total of fifteen British and French colonies remained in 1987. Figure 1 gives the cumulative number of decolonization events over time (dates listed in Appendix B).

The population-level time path of decolonization has the S-shaped curve characteristic of standard diffusion models. As discussed above, this implies that transition rates increase with time. A useful way to check this is to plot the integrated hazard, which is defined as

$$A(t) = \int_{t_0}^{t} r(s \mid t_0)ds \tag{13}$$

The slope of the integrated hazard thus gives the transition rate. The integrated hazard may be estimated nonparametrically following Aalen (1978) as

$$\hat{A}(t) = \sum_{j=t_1}^{t} d_j/r_j \tag{14}$$

where d_j is the number of events at t_j and r_j is the number of cases at risk at that time.

The increasing slope of the curve in Figure 2 shows that the rate is historically increasing. In the absence of complex models of true time dependence or temporally changing exogenous variables, this suggests the epidemic model as a plausible account of observed cluster-

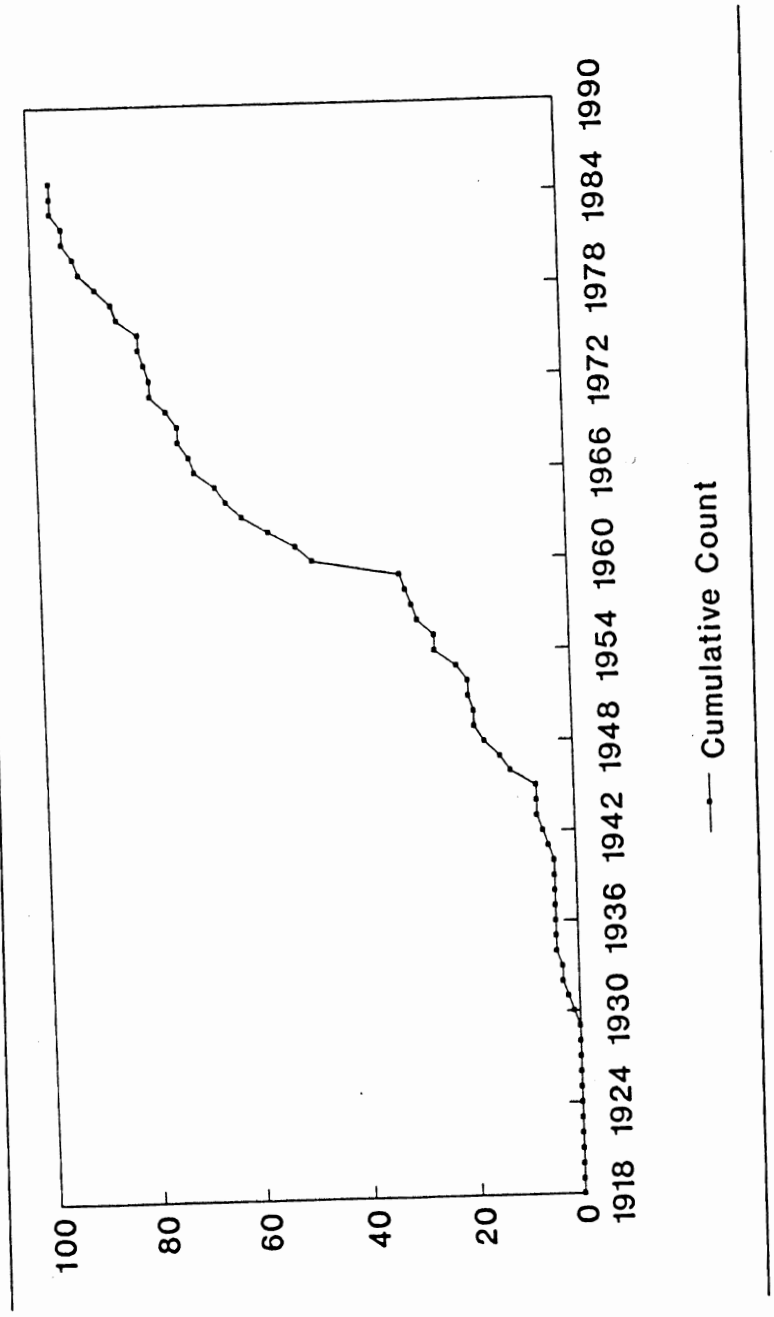

Figure 1: French and British Decolonization

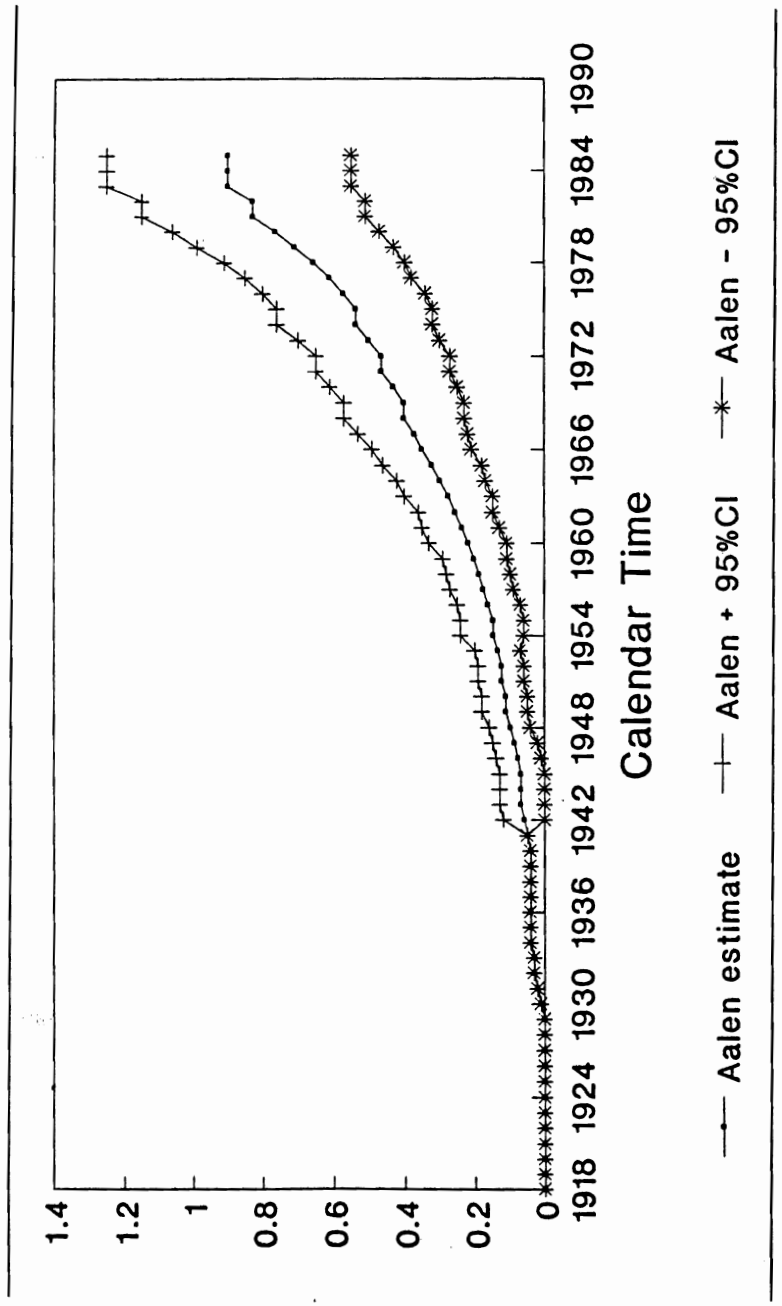

Figure 2: Integrated Hazard for Decolonization

ing. Unobserved heterogeneity would not produce this pattern and is only consistent with it if there are additional factors directly producing positive time dependence. Equally, diffusion along the lines of the salience model would not systematically produce an increasing rate over time.

CHANNELS OF DIFFUSION

Three social relations are investigated as possible channels of diffusion. The first is membership in the same empire. There are several reasons why diffusion might flow through empires more quickly than between them. For one, empires may form the critical reference groups for dependencies, so the independence of one colony sends a signal to other colonies subordinated to the same metropolitan power. In addition, the occurrence of decolonization within an empire may push the colonial power toward acceptance of decolonization for its other dependencies. A historical review shows that colonial powers often resist liberation movements vigorously at first: witness Britain in India, France in Indochina, the Netherlands in Indonesia (Grimal 1978). But the inability to maintain a key colony, particularly if the effort is militarily or politically costly, often seems to lead to a generalized acceptance of decolonization.

A second plausible channel of diffusion is region. This relation may also be thought to work through imitation in a reference group sense, with dependencies most aware of the events that occur to their neighbors. For example, the independence of the Gold Coast (as Ghana) in 1957 is often seen as catalyzing the development of nationalism in other African dependencies. In addition, regional linkages might channel diffusion through the material support provided by recently sovereign neighbors (for example, as staging grounds for guerrilla groups).

Finally, one might argue that decolonization anywhere should be important. The world systems and dependency literatures have argued that the globe is tightly connected, both politically and economically. For example, the United Nations and the World Court serve as forums where global ideologies can flower. The finding that diffusion is *not* structured by intra-imperial or intra-regional relations is thus substan-

TABLE 1: Descriptive Data on Decolonization by Empire and Region

Category	*Mean Date of Decolonization*	*Total Cases of Decolonization*	*Remaining Dependencies*
Empire			
British	1962.6	64	10
French	1957.5	32	5
Region			
North Africa and the Middle East	1956.0	18	0
Americas	1965.4	18	7
Sub-Saharan Africa	1961.7	34	2
Asia	1957.9	19	2
Oceania	1966.2	7	4

tively important. One benefit of examining the impact of particular social structures on diffusion is that it becomes possible to provide support for homogeneous mixing (though it is always possible that social structures unremarked by the analyst are operative).

The relations of interest here take the form of partitions of the population. Each dependency is assigned values identifying its empire (British of French) and its geographical region (the categories are the Americas, North Africa and the Middle East, sub-Saharan Africa, Asia, and Oceania). These and the dates of sovereignty are given in Appendix B. Table 1 gives some simple descriptive information on decolonization by empire and region.

RESULTS FOR EPIDEMIC MODELS

Table 2 displays results for five epidemic models of diffusion for British and French colonies between 1918 and 1987. Each of the exogenous variables measures the number of prior events within the dependency's social relation. Model A0 presents the results of a pure homogeneous mixing model. Models A1 and A2 analyze imperial and geographic linkages separately, comparing each to diffusion from dependencies outside the relation. Model A3 simultaneously considers the impact of the two relations as channels of diffusion. Model A4 adds counts of events outside both empire and region to compare these channels to the possibility of homogeneous mixing. All models are

TABLE 2: Epidemic Models of Transition Rates of Decolonization

Variable	*A0 Estimate (S.E.)*	*A1 Estimate (S.E.)*	*A2 Estimate (S.E.)*	*A3 Estimate (S.E.)*	*A4 Estimate (S.E.)*
α	−4.79** (.171)	−4.80** (.172)	−4.79** (.171)	−4.71** (.161)	−4.75** (.171)
Global	.028** (.002)				
Empire		.023* (.007)		.020** (.006)	.012 (.007)
Non-Empire		.034** (.008)			
Region			.076** (.011)	.072** (.013)	.070** (.012)
Non-Region			.016** (.004)		
Non-Empire and Non-Region					.020* (.010)
Likelihood Ratio χ^2 versus $r(t) = \exp(\alpha)$	86.2**	86.8**	103.4**	100.2**	103.4**
(*df*)	1	2	2	2	3
versus Model A0		.6	17.2**		
(*df*)		1	1		
versus Model A3					3.2
(*df*)					1

*$p < .05$; **$p < .01$.

estimated by the method of maximum likelihood using RATE (Tuma 1980). Likelihood ratio tests are presented comparing various nested models.

Model A0 presents the simplest diffusion model, where the variable "Global" counts the number of prior decolonization events anywhere in the population. The coefficient for this variable is significant, and the likelihood ratio test indicates that the model improves on a baseline model incorporating only a constant. This suggests that unstructured diffusion (diffusion under homogeneous mixing) provides a reasonable starting point for analysis of twentieth century British and French decolonization. In most studies of diffusion, this finding would also terminate the analysis. A great deal more can be learned by turning to the interplay of diffusion and social structure.

Model A1 indicates that the rate of decolonization is significantly increased by the number of previous decolonization events both within the empire and outside the empire. The notion that diffusion plays an important role in decolonization is again supported. But the important result is that the two coefficients have about the same magnitude. The likelihood ratio test comparing this model to Model A0 shows that numbers of events within and outside empires do not have significantly different effects.[8] British colonies do not appear more affected by independence events within the British Empire than they are by independence events in the French Empire (and vice-versa for the French). Imperial linkages thus do not seem to channel diffusion.

Regional linkages, on the other hand, have much stronger effects. While both effects are significant in Model A2, the coefficient for intra-region diffusion is about five times larger than the coefficient for inter-region diffusion, and the likelihood ratio test comparing the model to A0 shows this difference is statistically significant. The coefficient of .076 indicates that an additional previous decolonization event within a region multiplies the rate of decolonization by about 1.07. The independence of sixteen sub-Saharan dependencies in 1960 thus roughly triples the estimated decolonization rate in sub-Saharan Africa for 1961. In fact, an additional twelve sub-Saharan dependencies became sovereign in the next five years.

Model A3 directly contrasts the impact of imperial and regional linkages on the diffusion of independence. The results indicate that intra-regional linkages are more important than intra-imperial ones. Both coefficients are significant, but the effect for within-region diffusion is more than three times as large. Again, this relationship can be seen most vividly for Africa, where the majority of independence events in 1960 involved French colonies but the majority in the next few years involved British dependencies. The era of African decolonization succeeded earlier decolonization in Asia and the later independence of British colonies in the Caribbean.

Model A4 adds the effects of decolonization outside the dependency's empire and region to examine the evidence for homogeneous mixing. The addition of this variable does not significantly improve the fit of the model. Within-region events continue to exert

the largest effect on the rate of decolonization. In this analysis, their impact is more than three times larger than those of events within empires or events outside both empire and region. Neither of these two latter sources have significant effects.[9] By estimating the effects of the three sources of diffusion simultaneously, this formulation sums up the findings of the other models concisely. For British and French decolonization in the twentieth century, epidemic models indicate that the key social relation channeling diffusion is geographic proximity.

RESULTS FOR SALIENCE MODELS

Table 3 presents results for salience models of the diffusion of sovereignty. The first three models separately examine diffusion within empires, within regions, and from outside empire and region. The fourth model analyzes the rates of decay for imperial and regional channels of diffusion simultaneously. In these models, a strong impact of diffusion is suggested by a large difference between the rate at the time of the event and the asymptotic rate (the ratio $\exp(\beta)/\exp(\alpha)$ is large) and by a slow decline toward the asymptote (a small negative value for γ). Spells prior to first events within region and empire have been deleted in the same fashion for all models, permitting closer comparisons between them.

Model B1 examines intra-imperial diffusion. At the time of an event within a dependency's empire, the transition rate is .081 [exp(–4.01) + exp(–2.67]. It declines to about a quarter of this value, or .018 [exp(–4.01)], as time since the event becomes large. Return to this asymptotic value is quite rapid; one year after an event the rate is only about 2 percent larger than its asymptotic value.

Model B2 examines the impact of regional linkages. Time since the last event within the dependency's region has somewhat stronger effects. The rate at the point of an event is higher, equaling .159 (about eight times larger than the rate when t is very large). And although the rate of decline is about the same as for imperial diffusion, this larger initial shift means that a year after the event the rate is still about 34 percent larger than its aysmptotic value.

Decolonization occurring outside a dependency's empire and region appears to have no impact. In Model B3, β is barely significant

TABLE 3: Salience Models of Transition Rates of Decolonization

Variable	*B1 Estimate (S.E.)*	*B2 Estimate (S.E.)*	*B3 Estimate (S.E.)*	*B4 Estimate (S.E.)*
α	−4.01**	−3.95**	−3.31**	−4.16**
	(.21)	(.17)	(.10)	(.21)
β	−2.67**	−1.92**	−1.62*	−1.81**
	(.28)	(.31)	(.80)	(.27)
Rates of Decay γ				
Empire	−3.13*			−1.73*
	(1.45)			(.81)
Region		−3.09*		−.89
		(1.42)		(3.35)
Outside Empire and Region			6.52	
			(5.12)	
Likelihood Ratio χ^2				
versus $r(t) = \exp(\alpha)$	29.1**	51.0**	5.91	64.2**
(*df*)	2	2	2	3
versus Model B2				13.2**
(*df*)				1

*$p < .05$; **$p < .01$.

at a .05 level, and the rate of decay is very rapid. Most important, the model as a whole does not significantly improve over a constant rate model. The absence of an effect of diffusion from dependencies outside empire and region suggests that homogeneous mixing is not present.

Model B4 examines imperially and regionally channeled diffusion simultaneously, since these two social relations have significant effects when viewed in isolation. One is limited here to comparing rates of decay; the larger γ_k is in absolute value, the more quickly the impact of an event within that social relation wears off. The results again signal that intra-regional diffusion is considerably stronger than intra-imperial diffusion. The rate of decay for events within the dependency's empire is about double that of events within the region (−1.73 versus −.89). However, Model B4 does improve significantly over Model B2 by a likelihood ratio test, lending some support to the notion of intra-imperial diffusion.

For the problem of decolonization, Figures 1 and 2 suggest that a salience formulation is less appropriate than an epidemic formulation.

But overall, results for the salience and epidemic models are quite consistent. Both dismiss the hypothesis of homogeneous mixing, and both suggest that region is a more important channel of diffusion than empire.

CONCLUSIONS

The usual assumption of homogeneous mixing is sociologically uninformative. This article has sought to wed the analysis of diffusion to that of social structure. It has further argued for the simultaneous examination of competing notions of the structures relevant to diffusion. I would argue that such a strategy has the potential to enrich our substantive insight into the mechanisms and linkages underpinning diffusion.

Event history analysis is shown to offer a flexible approach to achieving these aims. Two diffusion models are presented. The first derives from an individual-level analogue of standard epidemic formulations; the second models the rate of adoption as a function of time since the last event. In both models, it is straightforward to allow rates to depend on prior adoptions by connected actors, rather than all members of the population. Traditional population level analyses of diffusion are unable to incorporate even the simplest forms of heterogeneity discussed here. And regression-based approaches are ill-suited to analyze temporal processes that event history methods are designed for, due to difficulties handling censored observations and keeping causes and effects in the correct time order.

To illustrate these approaches, twentieth-century decolonization in the British and French empires is modeled as a diffusion process. Both epidemic and salience models identify region as the key channel of diffusion and suggest that diffusion operating through empires or the world system as a whole is less important. These findings emphasize the ties linking dependencies to each other and downplay linkages between dependency and metropolis. The case of decolonization thus illustrates that adding social structure to diffusion models may not only produce a more fine-grained analysis of diffusion but also enlarge our understanding of social structure.

APPENDIX A

Illustrative reformatting of a hypothetical data set for diffusion analysis.

Dataset 1 (original data)

ID	ST	ET	SS	ES	G
1	1918	1934	1	2	A
2	1918	1945	1	2	A
3	1918	1951	1	1	A
4	1918	1973	1	2	A
5	1925	1952	1	2	B
6	1918	1960	1	1	B

Dataset 2 (reformatted data, where G defines subgroups within which diffusion is thought to occur)

ID	ST	ET	SS	ES	G	*n(t)*	ST_e	ET_e
1	1918	1934	1	2	A	0	—	—
2	1918	1934	1	1	A	0	—	—
2	1934	1945	1	2	A	1	0	11
3	1918	1934	1	1	A	0	—	—
3	1934	1945	1	1	A	1	0	11
3	1945	1951	1	1	A	2	0	6
4	1918	1934	1	1	A	0	—	—
4	1934	1945	1	1	A	1	0	11
4	1945	1973	1	2	A	2	0	18
5	1925	1952	1	2	B	0	—	—
6	1918	1952	1	1	B	0	—	—
6	1952	1960	1	1	B	1	0	8

ST : historical time at the start of the spell
ET : historical time at the end of the spell
SS : state occupied at the start of the spell
ES : state occupied at the end of the spell
G : hypothetical channel of diffusion
n(t) : number of prior events
ST_e : time since last event at the start of the spell
ET_e : time since last event at the end of the spell
— : undefined

APPENDIX B

Dates of sovereignty for British and French dependencies 1918-1987, with empire and region category.

DEPENDENCY	DATE	EMPIRE	REGION
Weihaiwei	10/ 1/ 1930	1	4
Canada	12/12/ 1931	1	1
Iraq	10/ 3/ 1932	1	2
Union of South Africa	/ / 1934	1	3
Levant States	/ / 1941	2	2
Australia	/ / 1942	1	5
Great Lebanon	12/27/ 1943	2	2
French Guiana	3/19/ 1946	2	1
Guadeloupe	3/19/ 1946	2	1
Martinique	3/19/ 1946	2	1
Reunion	3/19/ 1946	2	3
Transjordan	3/22/ 1946	1	2
Indian Princely States	/ / 1947	1	4
New Zealand	11/25/ 1947	1	5
Ceylon	2/ 4/ 1948	1	4
Palestine	5/14/ 1948	1	2
Burma	12/10/ 1948	1	4
Newfoundland	2/18/ 1949	1	1
Bhutan	8/ 8/ 1949	1	4
Oman	/ / 1951	1	2
Egypt	6/18/ 1953	1	2
Laos	10/15/ 1953	2	4
Cochin China	6/ 4/ 1954	2	4
Annam	7/21/ 1954	2	4
Cambodia	7/21/ 1954	2	4
Tonkin	7/21/ 1954	2	4
Anglo-Egyptian Sudan	1/ 1/ 1956	1	2
Morocco	3/ 2/ 1956	2	2
Tunis	3/20/ 1956	2	2
Gold Coast	2/ 7/ 1957	1	3
French Guinea	10/ 2/ 1958	2	3
Singapore	/ / 1959	1	4
Cameroun	1/ 1/ 1960	2	3

APPENDIX B continued

DEPENDENCY	DATE	EMPIRE	REGION
Togo	4/27/ 1960	2	3
Federation of Mali	6/20/ 1960	2	3
Madagascar	6/26/ 1960	2	3
Br. Somaliland	6/26/ 1960	1	3
Dahomey	8/ 1/ 1960	2	3
Niger	8/ 3/ 1960	2	3
Upper Volta	8/ 5/ 1960	2	3
Ivory Coast	8/ 7/ 1960	2	3
Chad	8/11/ 1960	2	3
Ubangi Shari	8/13/ 1960	2	3
Middle Congo	8/15/ 1960	2	3
Cyprus	8/16/ 1960	1	2
Gabon	8/17/ 1960	2	3
Federation of Nigeria	10/ 1/ 1960	1	3
Mauritania	11/28/ 1960	2	3
Sierra Leone	4/27/ 1961	1	3
Kuwait	6/19/ 1961	1	2
Tanganyika	12/ 9/ 1961	1	3
Algeria	7/ 3/ 1962	2	2
Jamaica	8/ 6/ 1962	1	1
French India	8/15/ 1962	2	4
Trinidad and Tobago	8/31/ 1962	1	1
Uganda	10/ 9/ 1962	1	3
Br. North Borneo	7/ 9/ 1963	1	4
Sarawak	7/ 9/ 1963	1	4
Malaya	9/16/ 1963	1	4
Zanzibar	12/ 9/ 1963	1	3
Kenya	12/12/ 1963	1	3
Nyasaland	7/ 6/ 1964	1	3
Malta	9/21/ 1964	1	2
N. Rhodesia	10/24/ 1964	1	3
The Gambia	2/18/ 1965	1	3
Maldives	7/26/ 1965	1	4
British Guiana	5/26/ 1966	1	1
Bechuanaland Botswana	9/30/ 1966	1	3

APPENDIX B continued

DEPENDENCY	DATE	EMPIRE	REGION
Basutoland	10/ 4/ 1966	1	3
Barbados	11/30/ 1966	1	1
F. of S. Arabia	11/30/ 1967	1	2
Mauritius	3/12/ 1968	1	3
Swaziland	9/ 6/ 1968	1	3
Tonga	6/ 4/ 1970	1	5
Fiji	10/10/ 1970	1	5
Bahrain	8/15/ 1971	1	2
Qatar	9/ 3/ 1971	1	2
Trucial States	12/ 2/ 1971	1	2
Bahamas	7/10/ 1973	1	1
Grenada	2/ 7/ 1974	1	1
Comoro Islands	1976	2	3
Br. Indian Ocean Territory	6/29/ 1976	1	4
Seychelles	6/29/ 1976	1	4
Ste. Pierre et Miquelon	7/15/ 1976	2	1
Djibouti	6/27/ 1977	2	3
Solomon Island	7/ 7/ 1978	1	5
Ellice Island	10/ 1/ 1978	1	5
Dominica	11/ 3/ 1978	1	1
St. Lucia	2/22/ 1979	1	1
Gilbert Island	7/12/ 1979	1	5
St. Vincent	10/27/ 1979	1	1
Southern Rhodesia	4/18/ 1980	1	3
Br. Honduras-Belize	9/21/ 1981	1	1
Antigua and Barbuda	11/ 1/ 1981	1	1
St. Kitts-Nevis	9/19/ 1983	1	1
Brunei	12/31/ 1983	1	4

EMPIRE: 1 = British, 2 = French
REGION: 1 = Americas, 2 = North Africa and Middle East, 3 = Sub-Saharan Africa, 4 = Asia, 5 = Oceania

NOTES

1. An alternative kind of analysis brings in social structure in a quite different way: Position within a social structure is a characteristic affecting individual likelihoods of adoption. See Cancian (1967) on the innovativeness of different socioeconomic classes.

2. γ is Euler's constant and $\phi(X)$ the digamma function.

3. Hannan and Freeman's (1987) study of organizational births also models transition rates as a function of the number of previous events, with a different rationale. They use $n(t)$ to measure the legitimacy of a new organizational form and $n(t)^2$ to capture the intensity of intra-population competition. Diekmann (1989) develops related models but inappropriately equates population- and individual-level expressions.

4. Analogously, one could also examine the effects of some appropriate measure of time t by modeling the rate as

$$r_i(t) = \exp[\alpha_i + \Sigma_{k=1}^{K} \beta_k n_{ik}(t)] + \exp(\gamma t)$$

5. Alternatively, one might assume that before first events $t_{ik} = \infty$ and set it to a very large value. It seems to me that such an approach would be sensitive to the large value chosen and potentially destabilize estimation without contributing much to the substantive concerns of the analysis.

6. For seven dependencies only the year was ascertained. This does not seem overly problematic, since five of these were the only decolonization events that year.

7. 1918 is the starting date for dependencies formed before that date. Dependencies formed after 1918 enter into the analysis in the year they are formed.

8. In these analyses, more general models often involve the relaxation of equality constraints, rather than constraints setting parameters to zero. Here, Model A1 generalizes Model A0 since A0 could be restated as A1 with the constraint that the impact of prior imperial and non-imperial decolonization events are equal. Comparison of A1 to A0 is thus asymptotically equivalent to a test for difference between the coefficients of intra- and inter-imperial diffusion.

9. Supplementary analyses added historical time, as discussed in Note 4. Regional diffusion again significantly increased the rate of decolonization, while intra-imperial and extra-imperial and regional events had no effect. A variety of measured covariates would be needed, however, to attempt to control for unobserved heterogeneity in the population.

REFERENCES

Aalen, Odd O. 1978. "Nonparametric Inference for a Family of Counting Processes." *Annals of Statistics* 6:701-26.

Bailey, Norman T. 1975. *The Mathematical Theory of Infectious Diseases and Its Applications*. New York: Hafner Press.

Bartholomew, David J. 1982. *Stochastic Models for Social Processes*. New York: Wiley.

Blumen, Isadore, Marvin Kogan, and Philip J. McCarthy. 1955. *The Industrial Mobility of Labor as a Probability Process*. Ithaca, NY: Cornell Studies in Industrial and Labor Relations.

Burt, Ronald S. 1987. "Social Contagion and Innovation: Cohesion Versus Structural Equivalence." *American Journal of Sociology* 92:1287-1334.

Cancian, Frank. 1967. "Stratification and Risk-Taking: A Theory Tested on Agricultural Innovation." *American Sociological Review* 32:912-27.

Coleman, James S. 1964. *Introduction to Mathematical Sociology*. London: Free Press.

Coleman, James S., Elihu Katz, and Herbert Menzel. 1966. *Medical Innovation*. New York: Bobbs-Merrill.

Conell, Carol and Samuel Cohn. 1989. "Environmental Variation and Diffusion in Strike Waves: French Coal Mining 1890-1935." Paper presented at the Annual Meeting of the American Sociological Association.

Diekmann, Andreas. 1989. "Diffusion and Survival Models for the Process of Entry into Marriage." Pp. 170-83, in *Event History Analysis in Life Course Research*, edited by K. U. Mayer and N. B. Tuma. Madison: University of Wisconsin Press.

Doreian, Patrick. 1981. "Estimating Linear Models in Spatially Distributed Data." Pp. 359-88, in *Sociological Methodology*, edited by S. Leinhardt. San Francisco: Jossey-Bass.

Grimal, Henri. 1978. *Decolonization*. London: Routledge & Kegan Paul.

Hannan, Michael T. and John Freeman. 1987. "The Ecology of Organizational Founding: American Labor Unions, 1836-1985." *American Journal of Sociology* 92:910-43.

Holland, Roy F. 1985. *European Decolonization 1918-1981: An Introductory Survey*. New York: St. Martin's.

Nogueira, Franco. 1963. *The United Nations and Portugal: A Study of Anti-Colonialism*. London: Sidgwick and Jackson.

Pitcher, Brian L., Robert L. Hamblin, and Jerry L. L. Miller. 1978. "The Diffusion of Collective Violence." *American Sociological Review* 43:23-35.

Spilerman, Seymour. 1970. "The Causes of Racial Disturbances: A Comparison of Alternative Explanations." *American Sociological Review* 35:627-49.

Tuma, Nancy B. 1980. *Invoking RATE*. Menlo Park, CA: Stanford Research International.

Tuma, Nancy B. and Michael T. Hannan. 1984. *Social Dynamics: Models and Methods*. Orlando, FL: Academic Press.

Zielske, Hubert A. and Walter A. Henry. 1980. "Remembering and Forgetting Television Ads." *Journal of Advertising Research* 20:7-13.

David Strang is an assistant professor of sociology at the University of Iowa. His ongoing research projects include analyses of decolonization, the expansion of the welfare state, and the population dynamics of health maintenance organizations. His substantive and methodological interests focus on institutional theory and their modeling as dynamic processes.