Published in final edited form as:

Sociol Methodol. 2009 August 1; 39(1): 185–232. doi:10.1111/j.1467-9531.2009.01212.x.

EFFECTS OF EXPOSURE ON PREVALENCE AND CUMULATIVE RELATIVE RISK: DIRECT AND INDIRECT EFFECTS IN A RECURSIVE HAZARD MODEL

Lawrence L. Wu and

Department of Sociology New York University

Steven P. Martin

Department of Sociology University of Maryland, College Park

Abstract

This paper outlines decomposition methods for assessing how exposure affects prevalence and cumulative relative risk. Let $\mathbf{x}$ denote a vector of exogenous covariates and suppose that a single dimension of time t governs two event processes T_1 and T_2 . If the occurrence of the event T_1 determines entry into the risk of the event T_2 , then subgroup variation in T_1 will affect the prevalence T_2 , even if subgroups in the population are otherwise identical. Although researchers often acknowledge this phenomenon, the literature has not provided procedures to assess the magnitude of an exposure effect of T_1 on the prevalence of T_2 . We derive decompositions that assess how variation in exposure generated by direct and indirect effects of the covariates $\mathbf{x}$ affect measures of absolute and relative prevalence of T_2 . We employ a parametric but highly flexible specification for baseline hazard for the T_1 and T_2 processes and use the resulting parametric proportional hazard model to illustrate the direct and indirect effects of family structure when T_1 is age at first sexual intercourse and T_2 is age at a premarital first birth for data on a cohort of nonhispanic white U.S. women.

Social scientists have long understood that exposure and prevalence are related—that how long people are exposed to the risk of an outcome will affect the prevalence of that outcome. Consider nonmarital births, which have been increasingly prevalent in the United States and which accounted for 38 percent of all U.S. births in 2006 (Hamilton, Martin, and Ventura 2007). Declining trends in the sexual activity of U.S. teens are often hailed as promising news in efforts to curb nonmarital childbearing among unmarried teen women; conversely, a declining age at first intercourse within particular teen subpopulations are often seen as cause for concern. Implicit is a classic social scientific insight—that all else being equal, later entry into sexual activity will decrease an unmarried young woman's exposure to the risk of a birth. An implication is that in aggregate populations, the prevalence of first births prior to a first marriage will then vary with shifts in the onset of sexual activity—that is, in exposure to risk.

The above example has several components that can be readily formalized. We have posited two events of interest, with the occurrence and timing of a first event, T_1 , placing individuals at risk of second event, T_2 . Variation in T_1 timing will generate variation in exposure to risk of the T_2 event, which in turn implies variation in the prevalence of the T_2 event.

Issues of exposure and prevalence also extend naturally to the influence of covariates. For example, a substantial literature has documented that nonmarital birth risks vary with social and demographic characteristics, as does women's age at onset of sexual activity. Covariates affecting both event processes will then have both direct and indirect effects on the second event process in ways that are similar to the classical recursive structural equation model for metric outcomes (see, e.g., Duncan 1975).

These issues also often arise in questions that are of considerable substantive and policy importance. Take, for example, the finding that women from single mother families have a higher probability of a premarital first birth than women who resided in families with two biological parents. Some of this greater prevalence will be due to increased exposure due to the earlier average onset of sexual activity among women residing in single mother families; likewise some of this greater prevalence will be due to higher premarital first risks for these women in the period following onset of sexual activity. A natural question then is which of these is larger. This paper answers this question by providing methods that decompose the effect of a covariate into direct and indirect components, including an indirect component reflecting the influence of exposure on prevalence.

Although there have long existed methods for tracing the implications of differential exposure to risk in aggregate populations (see, e.g., Kitagawa 1955; Das Gupta 1993; Smith, Morgan, and Koropecj-Cox 1996), we lack corresponding methods relating exposure and prevalence for the hazard regression context and when using individual-level data. This paper fills this gap for the case in which the occurrence of one event determines entry into risk of a second event process.

Focusing attention on situations in which a first event governs entry into the risk of a second event may at first glance appear highly restrictive, but such situations are common to a number of phenomena studied by social scientists. Examples include career mobility, in which a first job necessarily precedes a second job (Tuma 1976); movement through educational institutions that typically requires completion of earlier levels before proceeding to higher levels (Mare 1980); birth parity progressions, in which a third birth is necessarily preceded by two prior births (Westoff, Potter, and Sagi 1963); and marital transitions, in which divorce can occur only after marriage (Hannan, Tuma, and Groeneveld 1978).

A consistent pattern that emerges from our empirical decompositions is that the direct effects of covariates are, in general, larger in magnitude than their corresponding indirect effects. Although these findings will for many social scientists appear unsurprising—direct effects very often dominate indirect effects—they could also be taken to suggest a smaller role of exposure than is often presumed. But because our data are observational, it is not possible to regard these findings as identifying causal effects, but rather as providing estimates of the regression-adjusted association between various quantities. However, we also show below that unobservables must have highly specific characteristics if indirect effects are to dominate direct effects, a result that follows from the structure of our formal model. This implies a far smaller class of relevant unobserved confounds for these hypotheses than is usually the case.

The paper is organized as follows. We begin by stating the problem informally and comparing it with the classical recursive structural equation model. We then formalize ideas. We outline procedures for decomposing direct and indirect effects of covariates on the prevalence of the second event, which shows how variation in the timing of the first event affects prevalence of the second event. The essential ideas involve left truncating the hazard rate for the second event on the timing of the first event and decomposing the integrated hazard and survival functions for the second event into components representing the direct

effects of covariates and indirect effects of exposure. These formal results apply to both observed and unobserved covariates. We then illustrate these methods using data on age at first sexual intercourse and at a premarital first birth from the 1988 National Survey of Family Growth.

THEORY

We begin by comparing the classical three-variable recursive model for an uncensored metric outcome with an analogous three-variable recursive hazard model. We begin informally and then formalize details. To simplify the discussion, we consider two dichotomous events, T_1 and T_2 , in which there are two transitions of interest: an initial transition to T_1 , and a second transition from T_1 to T_2 . We then present the familiar derivation for the direct effect of exogenous covariates $\mathbf{x}$ on T_2 . A formally similar derivation is given for the effect of exposure on prevalence of T_2 resulting from variation in T_1 . The section concludes with derivations for the indirect effects of $\mathbf{x}$ on T_2 prevalence.

The classical recursive model for uncensored metric outcomes can be depicted as:

$$\begin{array}{ccc} \mathbf{x} & \xrightarrow{b_{yx}} & Y \\ & \searrow b_{zx} & \downarrow b_{zy} \\ & & Z \end{array}$$

or equivalently as the pair of equations

$$\begin{aligned} Y &= \mathbf{b}_{yx}\mathbf{x} + \epsilon_y \\ Z &= b_{zy}Y + \mathbf{b}_{zx}\mathbf{x} + \epsilon_z \end{aligned} \quad (1)$$

where Y is a metric variable that is predetermined with respect to a metric variable Z , and $\mathbf{x}$ is a vector of covariates that are predetermined with respect to both Y and Z .¹ A classic sociological example is status attainment (Featherman and Hauser 1978), in which Z is a measure of the respondent's occupational status during early adulthood, Y is the respondent's educational attainment, and $\mathbf{x}$ includes background factors such as number of siblings and the occupational status and educational attainment of the respondent's father.

Using (1), one can decompose the gross effect of covariate x into a direct effect b_{zx} and an indirect effect $b_{zy} \times b_{yx}$ of a covariate x on Z , which follows from:

$$E[Z] = b_{zy}(\mathbf{b}_{yx}\mathbf{x}) + \mathbf{b}_{zx}\mathbf{x}, \quad (2)$$

under the usual assumptions that $E[\epsilon_y] = E[\epsilon_z] = 0$. Alternatively, consider a homogeneous group of individuals and a hypothetical treatment in which the value of a covariate is shifted from x to $x + \Delta$ for some individuals (a "treatment" group) but not for some other individuals (a "control" group). Under (1), a "treatment" corresponding to $\Delta = 1$ will generate an increase in Y of b_{yx} , but Y will also affect Z as a right-hand-side covariate in the Z equation. From (2), we have the resulting indirect effect of x on Z equal to $b_{zy} \times b_{yx}$.²

¹"Predetermination" of Y with respect to Z is a central assumption in our model. Examples in which this assumption is violated include situations in which Y and Z are jointly determined.

²Our use of "treatment" and "control" is meant to clarify the logic underlying these comparisons, not to assert that causal effects can be obtained using such models on observational data. Similarly, our use of the terms "direct effect" and "indirect effect" should not be taken as implying that these quantities are causal.

Suppose instead that Z is an event that may be subject to censoring. If Y is a metric variable, then no particular complications arise in this “mixed” case, and much of the apparatus familiar from the more classical metric outcome setting carries over.

In this paper, we treat the case in which both Y and Z are events that may be subject to censoring. Under these restrictions, we derive direct and indirect effects of $\mathbf{x}$ on Z that are roughly analogous to those in the classical recursive model for a metric variable Y . There are three ways, however, in which this case varies from the classical one. A first observation is that the timing of the event Y determines entry into the risk of the event Z . That is, prior to the occurrence of the event Y , individuals are, by assumption, not at risk of the event Z ; hence, there can be no direct effect of $\mathbf{x}$ on Z prior to the occurrence of Y , an issue that does not arise in recursive models with a metric variable Y . Accounting for this difference involves conditioning in the likelihood for Z on entry into risk at time Y , where Y provides the so-called left-truncation time for the process Z .

A second observation is that variation in Y affects the duration that individuals are exposed to the risk of Z . Such variation will yield an effect of exposure on the *prevalence* of Z even in otherwise homogeneous populations. To see this informally, consider a population with two homogeneous subgroups, A and B , that differ only in the timing of Y . If the event Y occurs later in group A than in B , then the event Z will occur less frequently in group A than in B because members in group A spend less time exposed to the risk of event Z than members in group B . Because groups A and B are identical in all respects save for the timing of Y , differences in the prevalence of Z arise solely from differences in exposure.

The above discussion implies *two* indirect effects of $\mathbf{x}$ on the prevalence of Z . The first arises because the timing of Y determines entry into the risk of Z . The second is a hazard model variant of the usual indirect effect $\mathbf{b}_{yx}$ of $\mathbf{x}$ on Z for metric outcomes, in which (1) $\mathbf{x}$ affects the *risk* of Y , (2) Y risks determine the *timing* of Y , and (3) the *timing* of Y affects the prevalence of Z . Thus, the structure of this problem is more complicated than the classic recursive model for a metric Y and Z , in which there is a single indirect effect of $\mathbf{x}$ on Z .

Another difference between recursive models for metric and event history outcomes involves censoring. As noted above, observed differences in Y may be due to compositional differences as reflected in the covariates $\mathbf{x}$; this is true for recursive models in both settings. However, in a hazard model context, two additional complications arise: first, some individuals may not experience the event Y during the period of observation; and second, some individuals may never experience the event Y , even if observed for an infinitely long time. If the second condition holds, the distribution of Y is said to be defective and the prevalence of Z will necessarily be selected only on those who experience Y . A technical point about defective distributions is that the expected value for the timing of the event is undefined. To deal with this issue, we summarize distributional aspects of the timing of an event using percentiles such as the median rather than statistics such as the mean.

Finally, it should be noted that for metric outcomes, interpreting direct and indirect effects is straightforward because the indirect effect $b_{yx} \times b_{zy}$ is directly related to the regression coefficients b_{yx} and b_{zy} . But for the event history outcomes T_1 and T_2 , interpreting direct and indirect effects requires that these effects be translated into a common metric such as cumulative relative risk or prevalence.

To make clear the sequential dependence between Y and Z , we henceforth write them as (T_1, δ_1) and (T_2, δ_2) , respectively, where T_1 and T_2 are random variables denoting the event times and δ_1 and δ_2 are indicator variables for the two events. To simplify matters further, we often write T_1 and T_2 in place of (T_1, δ_1) and (T_2, δ_2) ; similarly, we engage in a slight abuse of language by referring, for example, to “the event T_1 ” or “the risk of T_2 .” We further

assume that both T_1 and T_2 are governed by a common dimension of time t , that all cases enter at risk of T_1 at $t = 0$, and that the occurrence of T_1 determines entry into the risk of the event T_2 .

To establish basic results, we proceed as follows. We begin with the transition to T_1 , reviewing standard results for the proportional hazard model on how variation in an exogenous covariate between two groups affects the predicted hazard rate for the transition to T_1 . We then trace how change in the predicted hazard rate for the transition to T_1 induces change in the distribution of T_1 event times. Percentiles of the T_1 event time distribution can then be used to show when a proportion p of persons with specified characteristics would be expected to experience this event. Inverting this problem shows how change in an exogenous covariate induces change in timing of the transition to T_1 . Thus, (1) change in an exogenous covariate x between two groups implies corresponding differences in the *rate* for the transition to T_1 , (2) group differences in the *rate* for the transition to T_1 imply differences in the T_1 distribution; and (3) group differences in *timing* of the transition to T_1 can be traced via percentiles of the T_1 distribution.

We then turn to the transition from T_1 to T_2 , again under a proportional specification. We first show how to model staggered entries of persons into the risk of the transition to T_2 using standard methods for left-truncated event history data. Covariate effects in such an equation can then be interpreted as direct effects for exogenous covariates. Because variation in T_1 induces variation in the exposure to risk of T_2 , we show how differences in the timing of T_1 affect the cumulative hazard of T_2 and two measures of T_2 prevalence; these relationships thus show how T_1 variation induces exposure consequences for the T_2 process. The final step combines this latter result with a previous result linking variation between groups in an exogenous covariate and variation in T_1 , with the resulting analytical expressions showing how covariates exert indirect effects on the T_2 process when persons become exposed to the risk of T_2 .

In the classical recursive model, the realization y of the random variable Y appears as a right-hand-side covariate in the Z equation. This can occur in an event history model as well, since one can condition on any aspect of an individual's past history when modeling the T_2 process, including, for example, an individual's age at time of entry into the left-truncated risk of T_2 . This would involve including the realization t_1 of T_1 as an ordinary right-hand-side covariate in the T_2 equation. The key point is that differences between two groups in an exogenous covariate will induce variation in T_1 , which in turn will induce variation in T_2 .

The Transition to T_1

Suppose that both T_1 and T_2 are governed by a common dimension of time t , that all cases enter at risk of T_1 at $t = 0$, and that the occurrence of T_1 determines entry into the risk of the event T_2 . Under these assumptions and further assuming a proportional hazard specification, the hazard rate for T_1 is given by:

$$r_1(t|\mathbf{x}) = q_1(t) \exp(\mathbf{a}\mathbf{x}), \quad (3)$$

where $q_1(t)$ denotes the baseline hazard for T_1 , $\mathbf{x}$ is a vector of covariates, and $\mathbf{a}$ denotes a vector of parameters to be estimated. Two fundamental quantities are the cumulative risk function $H_1(t|\mathbf{x})$ and the survivor function $S_1(t|\mathbf{x})$ given by:

$$H_1(t|\mathbf{x}) = \int_0^t r_1(u|\mathbf{x}) du = \exp(\mathbf{a}\mathbf{x}) Q_1(t) \quad (4)$$

where

$$Q_1(t) = \int_0^t q_1(s) ds, \quad (5)$$

and

$$S_1(t|\mathbf{x}) = \Pr(T_1 > t|\mathbf{x}) = \exp[-H_1(t|\mathbf{x})] = \exp[-\exp(\mathbf{a}\mathbf{x}) Q_1(t)]. \quad (6)$$

As in the linear regression context, compositional differences in the population will generate variation in T_1 . Note, however, that (6) provides predictions for a distribution of T_1 event times; thus unlike a linear regression, where a homogeneous subgroup will have a single predicted value of the outcome Y , under (6), a homogeneous subgroup will have a predicted distribution for the event times T_1 given by $S_1(t|\mathbf{x})$. Because of this, because some persons may be censored, and because the distribution of T_1 may be defective, it is more natural to compare percentiles of T_1 across groups than to compare expectations of T_1 across groups. Let t_{1p} denote the p th percentile of the distribution of T_1 . Suppose A and B are two groups of substantive interest, with covariate means given by $\bar{\mathbf{x}}_A$ and $\bar{\mathbf{x}}_B$, respectively. We wish to use (6) to obtain predictions for the timing of T_1 for groups A and B , evaluated at the p th percentile of the T_1 distribution. Since $S_1(t)$ varies between 0 and 1, consider the simple transformation from a p th percentile to the corresponding proportion given by $\pi = 1 - (p/100)$; then from (6), we have:

$$t_{1p} = Q_1^{-1}[-\log(\pi) / \exp(\mathbf{a}\mathbf{x})], \quad (7)$$

where $Q_1^{-1}(v)$ is the function such that if $v = Q_1(t)$ then $t = Q_1^{-1}(v)$. Given the above, note that compositional differences between groups A and B will generate differences in the p th percentiles of T_1 corresponding to:

$$\Delta t_{1p} = t_{1p}^B - t_{1p}^A = Q_1^{-1}\left[-\log(\pi) / \exp(\mathbf{a}\bar{\mathbf{x}}_B)\right] - Q_1^{-1}\left[-\log(\pi) / \exp(\mathbf{a}\bar{\mathbf{x}}_A)\right], \quad (8)$$

where t_{1p}^A and t_{1p}^B denote the p th percentiles of T_1 in groups A and B , respectively. Note that obtaining t_{1p}^A , t_{1p}^B , and Δt_{1p} requires inverting the function for the integrated hazard $Q(t)$ either analytically or numerically. Analytic expressions are available for choices of $q(t)$ such as the exponential, Weibull, and Gompertz models and piecewise variants of these models (see, e.g., self-citation 2 and Appendix 1); similarly, generalizing (7) and (8) to incorporate time-varying variables is straightforward (see Appendix 2). Note, however, that the Cox proportional hazard model (1972) is more difficult to use in this context because it does not specify a parametric form for $q(t)$; in particular, standard proposals for estimating the integrated hazard under a Cox model (see, e.g., Breslow 1974) will not, in general, yield a well-defined inverse function for $Q_1^{-1}(x)$.

The Transition from T_1 to T_2

As noted above, one way in which T_1 influences the process T_2 is that individuals are not at risk of T_2 until the occurrence of T_1 . In addition, when modeling T_2 , one can condition on any relevant aspect of an individual's history (Aalen 1978; Tuma and Hannan 1984), including the timing of the event T_1 . Consequently, one can treat the observed value t_1 as a

ordinary right-hand-side covariate in the same way as Y appears as a covariate in the linear regression of Z in the standard three-variable recursive model for metric outcomes. Let $u = t - t_1$ denote duration since sexual onset and let the hazard $r_2(t|t_1, \mathbf{x})$ for T_2 be given by:

$$r_2(t, u|t_1, \mathbf{x}) = q_{21}(t|t_1) q_{22}(u) \exp(\alpha t_1 + \mathbf{b}\mathbf{x}), \quad (9)$$

where $q_{21}(t|t_1)$ reflects the dependence of T_2 on t , given entry into risk at time t_1 , and $q_{22}(u)$ reflects the dependence of T_2 on duration u . To simplify the exposition below, we first focus attention on the simpler case in which $q_{22}(u) = 1$, deferring for the moment details of the more general case of age and duration dependence.

When $q_{22}(u) = 1$ in (9), the role of T_1 in determining entry into risk can be seen more easily in the expression for the survivor function, $S_2(t|t_1, \mathbf{x})$:

$$S_1(t|t_1, \mathbf{x}) = \exp\left[-\int_{t_1}^t r_2(s|t_1, \mathbf{x}) ds\right] = \exp\left[-\exp(\alpha t_1 + \mathbf{b}\mathbf{x}) \int_{t_1}^t q_{21}(s) ds\right], \quad (10)$$

where the quantity t_1 appears in (10) both as a right-hand-side covariate in the expression $\exp(\alpha t_1)$ and by left-truncating the period of risk via the lower limit of integration. As such, the T_1 equation in (6) provides, by assumption, the selection mechanism governing entry into risk of T_2 . If the distribution of T_1 is defective, then some fraction of the population will never experience the event T_1 and hence will never enter into risk of T_2 .

Direct effect of $\mathbf{x}$ on T_2 prevalence

The following derivation replicates an elementary result covered in standard texts on hazard methods (albeit with a slight modification for the left truncation by t_1). Suppose that groups A and B are identical in all respects except that x_1 for groups A and B differs by a constant, i.e., $x_{1B} = x_{1A} + \Delta$. Consider the cumulative relative risk defined as the ratio of the cumulative hazard for group B to that for group A . Under (9), this ratio is:

$$\begin{aligned} \frac{H_2(t|t_1, \mathbf{x}_B)}{H_2(t|t_1, \mathbf{x}_A)} &= \frac{\int_{t_1}^t q_{21}(s) \exp[\alpha t_1 + b_1(x_{1i} + \Delta) + \dots] ds}{\int_{t_1}^t q_{21}(s) \exp[\alpha t_1 + b_1 x_{1i} + \dots] ds} \\ &= \frac{\exp[\alpha t_1 + b_1(x_{1i} + \Delta) + \dots] \int_{t_1}^t q_{21}(s) ds}{\exp[\alpha t_1 + b_1 x_{1i} + \dots] \int_{t_1}^t q_{21}(s) ds} \\ &= \exp(b_1 \Delta). \end{aligned} \quad (11)$$

Thus, the direct effect on the cumulative relative risk of a shift from x_1 to $x_1 + \Delta$ is given by the usual estimate of relative risk.

How is (11) related to prevalence? Substantively, one might be interested in two quantities related to prevalence, one involving *absolute* prevalence—the arithmetic difference in prevalence between groups A and B —and the other involving *relative* prevalence—the ratio of prevalence for the two groups. Unlike the case for relative risk, the equations for the arithmetic and for the ratio difference do not simplify algebraically. Instead, one must solve equation 10 for each of the values of $\mathbf{x}$ to be compared, then subtract or divide the two prevalences as desired.

Effect of the timing of T_1 on T_2 prevalence

Variation in T_1 will affect the prevalence of T_2 even in otherwise homogeneous populations because some individuals will have longer durations of exposure to the risk of T_2 by virtue of quicker T_1 transitions. Standard hazard regressions *adjust* for such variations in exposure in the hazard rate, but do not quantify the magnitude of the effect of exposure on prevalence. However, such exposure effects of T_1 on T_2 can be derived via the same ideas as used above. Consider two groups of individuals, A and B , who are identical in all respects save for the timing T_1 , and suppose that the random variable T_{1A} is realized as t_1 for group A and that T_{1B} is realized as $t_1 + \Delta$ for group B . Then the cumulative relative risk is given by:

$$\begin{aligned} \frac{H_2(t|t_{1B}, \mathbf{x})}{H_2(t|t_{1A}, \mathbf{x})} &= \frac{\int_{t_1+\Delta}^t q_{21}(s) \exp[\alpha(t_1+\Delta) + \mathbf{b}\mathbf{x}] ds}{\int_{t_1}^t q_{21}(s) \exp[\alpha t_1 + \mathbf{b}\mathbf{x}] ds} \\ &= \frac{\exp[\alpha(t_1+\Delta) + \mathbf{b}\mathbf{x}] \int_{t_1+\Delta}^t q_{21}(s) ds}{\exp[\alpha t_1 + \mathbf{b}\mathbf{x}] \int_{t_1}^t q_{21}(s) ds} \\ &= \exp(\alpha\Delta) \left[\int_{t_1+\Delta}^t q_{21}(s) ds / \int_{t_1}^t q_{21}(s) ds \right]. \end{aligned} \quad (12)$$

To motivate the notion of an exposure effect, we postulated a counterfactual tracing the consequences for an otherwise homogeneous population at risk of T_2 in which the timing of T_1 was shifted, yielding a change in exposure to risk. Such a “pure” effect of exposure implies $\alpha = 0$ and involves only the bracketed ratio of integrals in (12). Thus, the “pure” effect of exposure generated by a shift from t_1 to $t_1 + \Delta$ is given by the bracketed ratio of integrals in (12).

By contrast, $\alpha \neq 0$ suggests that the observed realization t_1 of the random variable T_1 has an effect on T_2 as a usual covariate in the T_2 equation. Such a situation could arise if T_1 has an causal effect on T_2 conditional on the other covariates in the model, or if the realization t_1 of the random variable T_1 was correlated with unobserved covariates that influence T_2 . In our empirical application below, we rely on this second interpretation, interpreting α as reflecting the association of unobservables that are correlated with T_1 but that are not captured by the other covariates in (10).

Equation (12) provides a decomposition of the cumulative relative risk into two multiplicative components corresponding to the exposure effect given by the bracketed ratio of integrals and a more “standard” proportional effect of T_1 on T_2 given by $\exp(\alpha\Delta)$. Note that while the “standard” effect does not vary with t by assumption, the effect of exposure will in general vary in nonlinear ways with t . As a result, it can be useful to evaluate the effect of exposure over a range of t .

One objection to (9) is that t_1 enters linearly in the specification of $\log r_2(t)$. Weakening this assumption involves no major difficulties; for example, suppose that the effect of t_1 on $\log r_2(t)$ is not linear, but instead captured by some nonlinear function $f(t_1)$, i.e.,

$$r_2(t|t_1, \mathbf{x}) = q_{21}(t|t_1) \exp[f(t_1) + \mathbf{b}\mathbf{x}], \quad (13)$$

Employing (13) rather than (9) yields

$$\frac{H_2(t|t_{1B}, \mathbf{x})}{H_2(t|t_{1A}, \mathbf{x})} = \frac{\exp[f(t_1+\Delta)]}{\exp[f(t_1)]} \left[\int_{t_1+\Delta}^t q_{21}(s) ds / \int_{t_1}^t q_{21}(s) ds \right]. \quad (14)$$

Note in particular that the “standard” effect arising from the presence of t_1 as a right-hand-side covariate is more complicated in (14) than in (12), but that the bracketed term for exposure is identical in (12) and (14).

The expression in (12) concerns the cumulative relative risk but does not speak directly to prevalence. As noted above, one might assess prevalence by examining the difference in T_2 prevalence between groups B and A . This yields:

$$\begin{aligned} [1 - S_2(t|t_1, \mathbf{x}_B)] - [1 - S_2(t|t_1, \mathbf{x}_A)] &= (1 - \exp[-H_2(t|t_{1B}, \mathbf{x})]) - (1 - \exp[-H_2(t|t_{1A}, \mathbf{x})]) \\ &= \exp\left[-\exp(\alpha t_1 + \mathbf{b}\mathbf{x}) \int_{t_1}^t q_{21}(s) ds\right] - \\ &\quad \exp\left[-\exp(\alpha[t_1 + \Delta] + \mathbf{b}\mathbf{x}) \int_{t_1 + \Delta}^t q_{21}(s) ds\right]. \end{aligned} \quad (15)$$

Similarly, relative T_2 prevalence for groups B and A is given by the ratio:

$$\frac{1 - S_2(t|t_{1B}, \mathbf{x})}{1 - S_2(t|t_{1A}, \mathbf{x})} = \frac{1 - \exp[-H_2(t|t_{1B}, \mathbf{x})]}{1 - \exp[-H_2(t|t_{1A}, \mathbf{x})]} = \frac{1 - \exp\left[-\exp(\alpha[t_1 + \Delta] + \mathbf{b}\mathbf{x}) \int_{t_1 + \Delta}^t q_{21}(s) ds\right]}{1 - \exp\left[-\exp(\alpha t_1 + \mathbf{b}\mathbf{x}) \int_{t_1}^t q_{21}(s) ds\right]}. \quad (16)$$

Note that the decomposition in (14) of the indirect effect into standard and exposure effects does not hold for either (15) or (16).

Indirect effect of $\mathbf{x}$ on T_2 prevalence

Assessing the indirect effect of $\mathbf{x}$ on T_2 proceeds along the same formal lines as in the previous sections, with $\mathbf{x}$ affecting T_1 and T_1 affecting T_2 in two ways—via an indirect effect of exposure and a more “standard” indirect effect. A first step is to trace the effect of x on the timing of T_1 . Consider the pool of individuals who have not yet experienced the event T_1 and suppose that two groups, A and B , are identical in all respects save for their values of x_1 . As before, set $x_{1B} = x_{1A} + \Delta$; then from (8), the effect of composition on the timing of T_1 is given by:

$$\begin{aligned} \Delta t_{1p} &= t_{1Bp} - t_{1Ap} \\ &= Q_1^{-1}[-\log(\pi)/\exp(\mathbf{b}\mathbf{x}_B)] - Q_1^{-1}[-\log(\pi)/\exp(\mathbf{b}\mathbf{x}_A)] \\ &= Q_1^{-1}[-\log(\pi)/\exp(b_1(x_1 + \Delta) + \dots)] - Q_1^{-1}[-\log(\pi)/\exp(b_1 x_1 + \dots)], \end{aligned} \quad (17)$$

where $\pi = 1 - (p/100)$ and p corresponds to the p th percentile for the distribution of T_1 . Note that the function Q^{-1} is highly nonlinear; hence, the predicted effect of a covariate x on the percentile distribution T_1 will vary with the percentile p at which the effect is evaluated.

Recall from (12) that a shift from t_1 to $t_1 + \Delta$ influences the cumulative relative risk in two ways, through an indirect effect of exposure and a “standard” indirect effect. But a shift from x_1 to $x_1 + \Delta$ will induce a shift in t_1 , thus generating both direct and indirect effects for the T_2 equation. This is given by combining (12) and (17), from which one can derive the indirect effect of shifting x_1 to $x_1 + \Delta$ on the cumulative relative risk:

$$\frac{H_2(t|t_{1B}, \mathbf{x}_B)}{H_2(t|t_{1A}, \mathbf{x}_A)} = \exp(b_1 \Delta) \exp(\alpha \Delta t_{1p}) \left[\int_{t_{1A\pi} + \Delta t_{1p}}^t q_{21}(s) ds / \int_{t_{1A\pi}}^t q_{21}(s) ds \right] \quad (14)$$

Thus, the consequence of shifting from x_1 to $x_1 + \Delta$ in the T_1 equation appears in three places in the T_2 equation in (18): a direct effect of x_1 on T_2 represented by the quantity $b_1 \Delta$, and

two indirect effects of x_1 via T_1 —an indirect effect of exposure represented by the lower limit of integration in (18), and a more usual indirect effect represented by the quantity $\alpha\Delta t_{1p}$.

A shift from x_1 to $x_1 + \Delta$ will also generate differences between groups B and A in T_2 prevalence. As before, one may wish to assess these differences in either absolute or relative terms. The absolute difference in T_2 prevalence generated in this way is given by:

$$\begin{aligned} [1 - S_2(t|t_{1B}, \mathbf{x}_B)] - [1 - S_2(t|t_{1A}, \mathbf{x}_A)] \\ = \exp[-H_2(t|t_{1A}, \mathbf{x}_A)] - \exp[-H_2(t|t_{1B}, \mathbf{x}_B)] \\ = \exp\left[\exp[\alpha t_{1A\pi} + b_1 x_1 + \dots] \int_{t_{1A\pi}}^t q_{21}(s) ds\right] - \\ \exp\left[\exp[\alpha(t_{1A\pi} + \Delta t_{1p}) + b_1(x_1 + \Delta) + \dots] \int_{t_{1A\pi} + \Delta t_{1p}}^t q_{21}(s) ds\right]. \end{aligned} \quad (19)$$

Similarly, the relative difference in T_2 prevalence from shifting x_1 to $x_1 + \Delta$ is given by:

$$\begin{aligned} \frac{1 - S_2(t|t_{1B}, \mathbf{x}_B)}{1 - S_2(t|t_{1A}, \mathbf{x}_A)} &= \frac{1 - \exp[-H_2(t|t_{1B}, \mathbf{x}_B)]}{1 - \exp[-H_2(t|t_{1A}, \mathbf{x}_A)]} \\ &= \frac{1 - \left\{ \exp[\alpha(t_{1A\pi} + \Delta t_{1p}) + b_1(x_1 + \Delta) + \dots] \int_{t_{1A\pi} + \Delta t_{1p}}^t q_{21}(s) ds \right\}}{1 - \left\{ \exp[\alpha t_{1A\pi} + b_1 x_1 + \dots] \int_{t_{1A\pi}}^t q_{21}(s) ds \right\}}. \end{aligned} \quad (20)$$

Neither (19) nor (20) decompose into separable multiplicative components, as does (18) for the cumulative relative risk. However, there exist standard demographic procedures that let one scale quantities such that the direct and indirect components in (19) sum to the absolute difference in prevalence (Das Gupta 1993; see also Smith, Morgan, and Koropecj-Cox 1996). To motivate such a decomposition for the difference in absolute prevalence in (19), let

$$\begin{aligned} G_A(t) &= G(t|t_{1A}, \mathbf{x}_A) = 1 - S_2(t|t_{1A}, \mathbf{x}_A) \\ G_B(t) &= G(t|t_{1B}, \mathbf{x}_B) = 1 - S_2(t|t_{1B}, \mathbf{x}_B), \end{aligned} \quad (21)$$

where for notational clarity, we suppress the dependence of G on t and $\mathbf{x}$ by writing $G(t) = g(\gamma_1, \gamma_2, \gamma_3)$, that is, as a function of three terms—a direct effect γ_1 , an indirect effect of exposure γ_2 , and an indirect covariate effect γ_3 . Then roughly speaking, the main idea in such a Das-Gupta decomposition for (19) is to evaluate the influence of switching from group A to B across all possible permutations of γ_1 , γ_2 , and γ_3 . For example, one set of permutations will be:

$$\begin{aligned} g(\gamma_{1B}, \gamma_{2A}, \gamma_{3A}) &= 1 - \exp\left[-\exp(\alpha t_{1A} + \mathbf{b}\mathbf{x}_B) \int_{t_{1A}}^t q_{21}(s) ds\right] \\ g(\gamma_{1A}, \gamma_{2B}, \gamma_{3A}) &= 1 - \exp\left[-\exp(\alpha t_{1A} + \mathbf{b}\mathbf{x}_A) \int_{t_{1B}}^t q_{21}(s) ds\right] \\ g(\gamma_{1A}, \gamma_{2A}, \gamma_{3B}) &= 1 - \exp\left[-\exp(\alpha t_{1B} + \mathbf{b}\mathbf{x}_A) \int_{t_{1A}}^t q_{21}(s) ds\right] \end{aligned} \quad (22)$$

with similar expressions holding for the remaining possible permutations $g(\gamma_{1A}, \gamma_{2A}, \gamma_{3A})$, $\dots$, $g(\gamma_{1B}, \gamma_{2B}, \gamma_{3B})$.

To impose the standardization constraint that the sum of the three direct and indirect effects equal the total difference in (19), let Δg_1 , Δg_2 , and Δg_3 denote the standardized decomposition for the direct effect γ_1 , the indirect exposure effect γ_2 , and the indirect covariate effect γ_3 , respectively. Then consider the following three standardized decompositions:

$$\begin{aligned}\Delta g_1 &= g(\gamma_{1B}, \cdot, \cdot) - g(\gamma_{1A}, \cdot, \cdot) \\ &= \frac{2}{6} [g(\gamma_{1B}, \gamma_{2A}, \gamma_{3A}) - g(\gamma_{1A}, \gamma_{2A}, \gamma_{3A})] + \\ &\quad \frac{2}{6} [g(\gamma_{1B}, \gamma_{2B}, \gamma_{3B}) - g(\gamma_{1A}, \gamma_{2B}, \gamma_{3B})] + \\ &\quad \frac{1}{6} [g(\gamma_{1B}, \gamma_{2A}, \gamma_{3B}) - g(\gamma_{1A}, \gamma_{2A}, \gamma_{3B})] + \\ &\quad \frac{1}{6} [g(\gamma_{1B}, \gamma_{2B}, \gamma_{3A}) - g(\gamma_{1A}, \gamma_{2B}, \gamma_{3A})]\end{aligned}\quad (23)$$

$$\begin{aligned}\Delta g_2 &= g(\cdot, \gamma_{1B}, \cdot) - g(\cdot, \gamma_{1A}, \cdot) \\ &= \frac{2}{6} [g(\gamma_{1A}, \gamma_{2B}, \gamma_{3A}) - g(\gamma_{1A}, \gamma_{2A}, \gamma_{3A})] + \\ &\quad \frac{2}{6} [g(\gamma_{1B}, \gamma_{2B}, \gamma_{3B}) - g(\gamma_{1B}, \gamma_{2A}, \gamma_{3B})] + \\ &\quad \frac{1}{6} [g(\gamma_{1A}, \gamma_{2B}, \gamma_{3B}) - g(\gamma_{1A}, \gamma_{2A}, \gamma_{3B})] + \\ &\quad \frac{1}{6} [g(\gamma_{1B}, \gamma_{2B}, \gamma_{3A}) - g(\gamma_{1B}, \gamma_{2A}, \gamma_{3A})]\end{aligned}\quad (24)$$

$$\begin{aligned}\Delta g_3 &= g(\cdot, \cdot, \gamma_{1B}) - g(\cdot, \cdot, \gamma_{1A}) \\ &= \frac{2}{6} [g(\gamma_{1A}, \gamma_{2A}, \gamma_{3B}) - g(\gamma_{1A}, \gamma_{2A}, \gamma_{3A})] + \\ &\quad \frac{2}{6} [g(\gamma_{1B}, \gamma_{2B}, \gamma_{3B}) - g(\gamma_{1B}, \gamma_{2B}, \gamma_{3A})] + \\ &\quad \frac{1}{6} [g(\gamma_{1A}, \gamma_{2B}, \gamma_{3B}) - g(\gamma_{1A}, \gamma_{2B}, \gamma_{3A})] + \\ &\quad \frac{1}{6} [g(\gamma_{1B}, \gamma_{2A}, \gamma_{3B}) - g(\gamma_{1B}, \gamma_{2A}, \gamma_{3A})],\end{aligned}\quad (25)$$

where the three terms in (23)–(25) sum to the “total” difference:

$$\Delta g_1 + \Delta g_2 + \Delta g_3 = g(\gamma_{1B}, \gamma_{2B}, \gamma_{3B}) - g(\gamma_{1A}, \gamma_{2A}, \gamma_{3A}). \quad (26)$$

The expressions in (23)–(25) provide a Das Gupta-type decomposition for the absolute difference in (19). To obtain a similar decomposition for (20), rewrite the left-hand-side ratio in (20) as

$$\frac{1 - S_2(t|t_{1B}, \mathbf{x}_B)}{1 - S_2(t|t_{1A}, \mathbf{x}_A)} = \frac{g_B(\gamma_1, \gamma_2, \gamma_3)}{g_A(\gamma_1, \gamma_2, \gamma_3)}. \quad (27)$$

Then taking logarithms yields

$$\log \left(\frac{1 - S_2(t|t_{1B}, \mathbf{x}_B)}{1 - S_2(t|t_{1A}, \mathbf{x}_A)} \right) = \log g_B(\gamma_1, \gamma_2, \gamma_3) - \log g_A(\gamma_1, \gamma_2, \gamma_3). \quad (28)$$

But because (28) is an arithmetic difference in $\log g$, replacing g by $\log g$ in Eqs. (23)–(26) will providing an analogous decomposition for relative prevalence.

Age and duration dependencies in T_2 prevalence

We now turn to the more general case in which the transition from T_1 to T_2 varies both with age and duration. We generalize (9) by considering the case in which the hazard $r_2(t|t_1, \mathbf{x})$ for T_2 is given by a specification in which the age and duration dependence are separable:

$$r_2(t, u|t_1, \mathbf{x}) = q_{21}(t|t_1) q_{22}(u) \exp(\alpha t_1 + \mathbf{b}\mathbf{x}), \quad (29)$$

where $q_{21}(t|t_1)$ reflects the dependence of T_2 on t given entry into risk at time t_1 and $q_{22}(u)$ the dependence of T_2 on duration u , with $u = t - t_1$ denoting duration since T_1 onset. Under (29), the survivor function, $S_2(t, u|t_1, \mathbf{x})$ will be given by the double integral

$$\begin{aligned} S_2(t|t_1, \mathbf{x}) &= \exp \left[- \int_{t_1}^t \int_0^u r_2(s, v|t_1, \mathbf{x}) ds dv \right] \\ &= \exp \left[- \exp(\alpha t_1 + \mathbf{b}\mathbf{x}) \int_{t_1}^t q_{21}(s) ds \int_0^u q_{22}(v) dv \right]. \end{aligned} \quad (30)$$

As before, consider a thought experiment in which groups A and B are identical in all respects, including the timing of T_1 onset, except that x_1 for groups A and B differs by a constant, i.e., $x_{1B} = x_{1A} + \Delta$. Then from (30), the expression in (11) for the cumulative relative risk for group B to that for group A remains unchanged, since

$$\begin{aligned} \frac{H_2(t, u|t_1, \mathbf{x}_B)}{H_2(t, u|t_1, \mathbf{x}_A)} &= \frac{\int_{t_1}^t \int_0^u q_{21}(s) q_{22}(v) \exp[\alpha t_1 + b_1(x_{1i} + \Delta) + \dots] ds dv}{\int_{t_1}^t \int_0^u q_{21}(s) q_{22}(v) \exp[\alpha t_1 + b_1 x_{1i} + \dots] ds dv} \\ &= \exp(b_1 \Delta). \end{aligned} \quad (31)$$

By contrast, the absolute and relative difference in T_2 prevalence in (12) and (13) are now given, respectively, by

$$S_2(t, u|t_1, \mathbf{x}_A) - S_2(t, u|t_1, \mathbf{x}_B) = \exp[-H_2(t, u|t_1, \mathbf{x}_A)] - \exp[-\exp(b_1 \Delta) H_2(t, u|t_1, \mathbf{x}_A)], \quad (32)$$

and

$$\frac{1 - S_2(t, u|t_1, \mathbf{x}_B)}{1 - S_2(t, u|t_1, \mathbf{x}_A)} = \frac{1 - \exp[-H_2(t, u|t_1, \mathbf{x}_B)]}{1 - \exp[-H_2(t, u|t_1, \mathbf{x}_A)]} = \frac{1 - \exp[-\exp(b_1 \Delta) H_2(t, u|t_1, \mathbf{x}_A)]}{1 - \exp[-H_2(t, u|t_1, \mathbf{x}_A)]}. \quad (33)$$

Now consider two groups of individuals, A and B , who are identical in all respects save for the timing T_1 , and suppose that the random variable T_{1A} is realized as t_1 for group A and that T_{1B} is realized as $t_1 + \Delta$ for group B . Then the cumulative relative risk in (12) will now be given by the double integral:

$$\begin{aligned} \frac{H_2(t, u|t_{1B}, \mathbf{x})}{H_2(t, u|t_{1A}, \mathbf{x})} &= \frac{\int_{t_1 + \Delta}^t \int_0^{u - \Delta} q_{21}(s) q_{22}(v) \exp[\alpha(t_1 + \Delta) + \mathbf{b}\mathbf{x}] ds dv}{\int_{t_1}^t \int_0^u q_{21}(s) q_{22}(v) \exp[\alpha t_1 + \mathbf{b}\mathbf{x}] ds dv} \\ &= \exp(\alpha \Delta) \frac{\int_{t_1 + \Delta}^t q_{21}(s) ds \int_0^{u - \Delta} q_{22}(v) dv}{\int_{t_1}^t q_{21}(s) ds \int_0^u q_{22}(v) dv}. \end{aligned} \quad (34)$$

The difference in T_2 prevalence between groups B and A also involves a change to double integrals

$$\begin{aligned} S_2(t, u|t_1, \mathbf{x}_B) - S_2(t, u|t_1, \mathbf{x}_A) &= (1 - \exp[-H_2(t, u|t_1, \mathbf{x})]) - (1 - \exp[-H_2(t, u|t_1, \mathbf{x})]) \\ &= \exp\left[-\exp(\alpha[t_1 + \Delta] + \mathbf{b}\mathbf{x}) \int_{t_1 + \Delta}^t q_{21}(s) ds \int_0^{u - \Delta} q_{22}(v) dv\right] - \\ &\quad \exp\left[-\exp(\alpha t_1 + \mathbf{b}\mathbf{x}) \int_{t_1}^t q_{21}(s) ds \int_0^u q_{22}(v) dv\right], \end{aligned} \quad (35)$$

with the relative T_2 prevalence for groups B and A given by the ratio:

$$\frac{1 - S_2(t, u|t_1, \mathbf{x}_B)}{1 - S_2(t, u|t_1, \mathbf{x}_A)} = \frac{1 - \exp\left[-\exp(\alpha[t_1 + \Delta] + \mathbf{b}\mathbf{x}) \int_{t_1 + \Delta}^t q_{21}(s) ds \int_0^{u - \Delta} q_{22}(v) dv\right]}{1 - \exp\left[-\exp(\alpha t_1 + \mathbf{b}\mathbf{x}) \int_{t_1}^t q_{21}(s) ds \int_0^u q_{22}(v) dv\right]}. \quad (36)$$

Finally, we consider the general case in which a shift from x_1 to $x_1 + \Delta$ induces a shift from t_1 to $t_1 + t_{1p}$, thus generating both direct and indirect effects for the T_2 equation. Let $u = t - \Delta t_{1p}$; then the cumulative relative risk is given by combining (34) and (36):

$$\frac{H_2(t|t_1, \mathbf{x}_B)}{H_2(t|t_1, \mathbf{x}_A)} = \exp(b_1 \Delta) \exp(\alpha \Delta t_{1p}) \frac{\int_{t_{1A} + \Delta t_{1p}}^t q_{21}(s) ds \int_0^{u - \Delta t_{1p}} q_{22}(v) dv}{\int_{t_{1A}}^t q_{21}(s) ds \int_0^u q_{22}(v) dv}. \quad (37)$$

Thus as in (18), the consequence of shifting from x_1 to $x_1 + \Delta$ in the T_1 equation appears in three places in (36) and (37): a direct effect of x_1 on T_2 represented by the quantity $b_1 \Delta$, and two indirect effects of x_1 via T_1 —an indirect effect of exposure represented by the lower and upper limits of integration, and the more usual indirect effect represented by the quantity $\alpha \Delta t_{1p}$.

DATA

We illustrate these methods using data on white women from the 1988 National Survey of Family Growth (NSFG), a retrospective survey of women aged 15–44 in 1988. We examine two young adult transitions, initiation of sexual activity and a premarital first birth—that is, a first birth that occurs prior to a first marriage. NSFG respondents were asked to supply the calendar year and month of first sexual intercourse; we converted these data into age (in months) at first intercourse. For this transition, we censored a woman's data either at her age at the 1988 NSFG interview if she reported never having experienced sexual intercourse or at her age of first marriage if she reported initiating sexual activity during the same month of first marriage or at some later time. We constructed age at a premarital first birth using the NSFG first birth and first marriage histories. For this transition, we censored a woman's data either at her age at first marriage if she married prior to a first birth or at her age at interview if neither a birth nor marriage occurred prior to interview.

Our empirical examples use a uniform set of covariates to model both the T_1 transition—the transition to first sexual intercourse—and the T_2 transition—to a premarital first birth. The covariates we examine are: snapshot measures of the respondent's family structure at age 14, education of the respondent's mother, age of the respondent's mother at first birth, a dummy variable for Catholic religion at interview, and a dummy variable equal to one if we employed a hot-deck procedure to impute calendar month at first intercourse. Our family structure variables are snapshot measures that contrast women who resided in two-

biological, mother-only, step, and other types of families at age 14. As noted in the previous section, in modeling T_2 , one can also include the timing of T_1 as non-time-varying, and duration since T_1 as a time-varying, right-hand-side covariate. Age at first intercourse and duration since first intercourse, then, are the only variables that appear in the T_2 equation but not the T_1 equation.

The NSFG contains data on 4,911 white women. We dropped cases if a first birth was reported prior to first intercourse or if data on first intercourse, first birth, first marriage, or family structure were missing. Of the full sample of 4,911 white women, 157 were missing data on either first intercourse or a premarital first birth, and another 2 women were missing data on family structure, yielding a sample of 4,752 white women.

RESULTS

Figure 1 presents smoothed nonparametric estimates using a procedure described in self-citation 3 for the age-graded risk of entry into sexual activity, the age-graded risk of a premarital first birth, and the duration-graded risk of a premarital first birth conditional on entry into sexual activity. The first panel of Figure 1 plots smoothed nonparametric estimates of the logarithm of the hazard rate of first sexual intercourse by age, the middle panel plots two different estimates of the logarithm of the hazard rate for a premarital first birth, and the bottom panel plots estimates of the logarithm of the hazard rate for a premarital first birth by duration since sexual onset. In the upper two panels, the curves for the logarithm of the rate rise in a roughly linear fashion to about age 18.5, after which the curves decline, again in a roughly linear fashion.

In the middle panel of Figure 1, the two curves differ in the assumptions they make about when women become at risk of a premarital first birth. The solid curve presents estimates that do not place a woman at risk of a premarital first birth until she reports becoming sexually active; hence, for this curve, we use a woman's report of age at first intercourse to left-truncate her premarital birth history. The dotted curve presents estimates that ignore this left truncation; hence, while this curve can be viewed as the average of the logarithm of premarital first birth risks in the population, it ignores variation in onset of sexual activity and implicitly assumes that women are at risk of a premarital first birth even if they have not initiated sexual activity, an implausible assumption. A comparison of the two curves in the lower panel of Figure 1 shows that left truncation affects estimates substantially, with the curve ignoring left truncation systematically underestimating premarital first birth risks relative to the curve that incorporates left truncation. Differences between these two curves are especially apparent at younger ages; substantively, this reflects the tendency for premarital births risks to be especially high among teenaged women in the period just following the initiation of sexual activity.

The nonparametric estimates in the bottom panel of Figure 1 exhibit less variation in premarital first birth risks by duration, as opposed to the clear patterns of age variation observed in the middle panel of Figure 1. Based on these nonparametric results, we model age dependence in both the T_1 and T_2 equations using a splined piecewise Gompertz specification with nodes at ages 16.5, 18.5, and 21.0 (e.g. self-citation 4; Lillard 1993). For the T_2 equation, age at first sex was modeled as a simple fixed covariate, and duration-dependence (since onset of sexual activity) was modeled using a piecewise exponential specification for durations 0 to 35, 36 to 71, and 72 or more months.³

³An alternative to our piecewise splined Gompertz model is a piecewise constant model, which is commonly used by researchers to approximate a Cox model. Note that the piecewise constant model is a special case of a piecewise Gompertz model, since the latter yields a piecewise linear baseline for $\log r(t)$.

We now proceed to our parametric models. To simplify the discussion, we focus attention on tracing the direct and indirect effects of family structure on our two transitions; however, our underlying models control both for family structure and the other background characteristics. Finally, and as noted above, when modeling T_2 , one can also specify an effect for the observed timing of T_1 as a non-time-varying right-hand-side covariate.

Table 1 presents coefficient estimates for family structure and, in some models of premarital first birth risks, coefficient estimates for age at first intercourse and for duration since sexual onset.⁴

Model 1 presents estimates for the transition to first sexual intercourse. As in previous research, we observe higher risks for women who resided at age 14 in families that did not contain both biological parents relative to women who resided with both biological parents at age 14. The coefficient of .51 for the effect of residing in a mother-only family at age 14 indicates an age-specific rate of first sexual intercourse that is $\exp(.51)$ or 1.67 times higher than the rate for women who resided with both biological parents at age 14. For women residing in a step family at age 14 and for women in some other family arrangement at age 14, the model coefficients of .69 and .58 indicate rates of first sexual intercourse 1.99 and 1.79 times greater, respectively, than the rate for women who resided with both biological parents at age 14.

The remaining models in Table 1 examine the transition to a premarital first birth and reflect different assumptions about the effects of age at first intercourse on premarital first births and about the variation with duration in premarital first birth risks. Model 2 neither left-truncates exposure to risk using the respondent's reported age at first intercourse nor includes an effect for age at first intercourse specified as a non-time-varying right-hand-side covariate. Thus, Model 2 contains no term involving T_1 ; hence, it assumes that risks are *identical* (conditional on family structure and the other background variables) for women who have and have not initiated sexual activity. Model 3 differs from Model 2 by including the respondent's reported age at first intercourse as a non-time-varying right-hand-side covariate for the respondent's risk of a premarital first birth. Model 3 does condition on T_1 , but does so by including the realized value of T_1 for a given respondent as a non-time-varying right-hand-side covariate. Hence, it assumes, somewhat implausibly, that women are at risk of a premarital first birth both before and after initiation of sexual activity, but models variation in premarital first birth risks for women who initiate sexual activity at different ages by a proportionality constant. See also Kiernan and Hobcraft (1997).

By contrast, Models 4–6 use a woman's reported age at first sexual intercourse to left truncate her risk of a premarital first birth; thus, these models do not place a woman at risk of a premarital first birth until she reports initiating sexual activity. Model 5 differs from Model 4 by adding age at first sexual intercourse as a non-time-varying right-hand-side covariate. Model 6 uses the full specification in (9), incorporating one baseline hazard for age dependence and a second baseline hazard for duration dependence.⁵ Comparing coefficient estimates in Models 5 and 6 shows that despite substantially lower risks ($\exp(-.40) = .67$ and $\exp(-1.01) = .36$, corresponding to a 32 and 64 percent lower risk) at durations 32–71 and 72+ months since sexual onset, adding a baseline hazard for duration dependence in premarital first birth risks changes the coefficient estimates for the other covariates only slightly.

⁴Parameter estimates for all variables, as well as those for the baseline hazard for both transitions, are available upon request. As a sensitivity check, we also estimated covariate effects in Table 1 using a Cox specification; see Appendix Table 1. Estimated coefficients from the two specifications differ in only slight ways for nearly all covariates.

Table 2 gives predicted values for the median age at first intercourse by family structure at age 14 using estimates from Model 1 in Table 2, with estimates evaluated at the mean values for all background covariates. The predicted median age at first intercourse is 226.5, 213.0, 209.1, and 211.3 months for white women who resided at age 14 in a two biological parent family, a mother-only family, a step family, or in the residual category for other types of families, respectively. The second column gives deviations from the predicted median for women who resided in an intact family at age 14. The median age at first intercourse is between 13.4 months and 17.4 months later for women who resided in an intact family at age 14 relative to the other three types of nonintact families, with the magnitude of these deviations corresponding to the hazard model parameter estimates reported in column 1 of Table 1.

Table 3 evaluates the indirect effects of the family structure variables on the cumulative relative risk of a premarital first birth given by (18) for white women, using estimated coefficients from Model 6 of Table 1. For most women in this sample, exposure to the risk of a premarital birth is ended by a transition to first marriage, so time t in columns two and four can be viewed counterfactually as the risks that a hypothetical woman would experience following onset of sexual activity but prior to entry into marriage. As in the hazard regressions in Table 1, white women who resided in a two biological parent family at age 14 form the baseline group, while as in Table 2, we set the remaining background covariates to their respective sample means and assess variation in age at first intercourse using the predicted median age at onset.

We begin with the comparison in row 1 of Table 3. This row gives the indirect effect of a covariate x on T_2 , in this case, having resided in a mother-only family at age 14 on the cumulative relative risk of a premarital first birth, given in expression (18). Recall that the indirect effect of x on T_2 occurs through its influence on T_1 , which we evaluate using the predicted median of the T_1 distribution. From Table 2, we have that the predicted median age at first intercourse was 226.5 and 213.0 months for white women who resided in an intact and mother-only family at age 14, respectively; these in turn imply values of $t_1 = 226.5$ and $\Delta = -13.5$ in the first row of Table 3. In addition, the expression for the cumulative relative risk in (18) requires that one choose some age t at which to evaluate the cumulative relative risk. In the first row of results, we have chosen an age equal to 60 months after the predicted median age at onset for women in the omitted category—those who resided in a two-biological family at age 14; this corresponds to $t_1 + 60 = 286.5$ months, or roughly age 24.

What is the indirect effect of having resided in a mother-only family at age 14 on a premarital first birth, relative to having resided in an intact family at age 14? The total indirect effect, evaluated at age 286.5 months, is to raise the cumulative risk by 51.8 percent relative to women who resided in an intact family at age 14. This 51.8 percent increase can be decomposed into an indirect component for exposure, corresponding to 19.0 percent increase in the cumulative relative risk, and a more usual indirect component, corresponding to a 27.5 percent increase, with the decomposition $1.518 = 1.190 \times 1.275$ given by the

⁵In somewhat more detail, consider a woman who initiated sexual activity at age 20. Models 2 and 3 assume that the woman has a nonzero risk of a premarital first birth at all ages, both before and after age 20. These two models differ in that Model 3 adds a proportionality term α to Model 2; hence, these models assume that the risk of a premarital birth for a woman who initiated sexual activity at age 20 relative, say, to an otherwise identical woman who initiated activity at age 19, is $\exp(\alpha \times 20) / \exp(\alpha \times 19) = \exp[\alpha(20 - 19)] = \exp(\alpha)$, with this proportionality factor assumed constant across all ages for the woman, including those prior to her onset of sexual activity. By contrast, Models 4 and 5 place the woman at risk of a premarital first birth only after age 20; put another way, Models 4 and 5 assume that the woman's risk of a premarital first birth is identically zero prior to age 20. Then considering two women who initiated sexual activity at ages 19 and 20 but who are otherwise identical in all respects, Model 4 assumes that these two women have identical premarital birth risks after age 20, while Model 5 lets the risks differ for these two women after age 20 by the proportional factor $\exp(\alpha)$.

expression in (18). Put another way, these results state that, all else being equal, having resided in a mother-only family at age 14 is associated with a 13.5 month earlier entry into sexual activity relative to women who resided in an intact family at age 14. The consequence of this 13.5 month difference is to increase the cumulative relative risk of a premarital first birth by 51.8 percent over a 5 year period, with .37 of this increase (19.0/51.8) attributable to the 13.5 months of increased exposure to risk.

The next two rows of Table 3 present results for having resided in a step family or the residual “other” category of family situation at age 14. In both cases, the indirect effect of having resided in these nonintact family situations is to raise the cumulative risk of a premarital first birth relative to having resided in an intact family. Across all three rows, the indirect effect of exposure is roughly two-thirds the magnitude of the more usual indirect effect. Note that risks are highest for women who, at age 14, resided in a step family, followed by those who resided in other types of families, then for those who resided in a mother-only family. Because Table 3 reports estimates for the *indirect* effects of family structure on the cumulative relative risk of T_2 , this ordering of effects is generated by the parameter estimates for family structure in the T_1 equation (column 1 of Table 1), not for the parameter estimates for family structure in the T_2 equation (column 5 of Table 1).

The results presented in the remaining rows show that the magnitude of the indirect effects of exposure varies with the time frame used to gauge exposure. While the comparisons in rows 1–3 suppose that women remain at risk of a premarital first birth during the first 60 months after the median age of first sex for women from an intact family, the comparisons in rows 4–6 extend the exposure time to 90 months. As noted earlier, women typically exit the risk of a premarital first birth through first marriage, and the 90 month exposure time will not be representative of the average cumulative exposure time for respondents in the 1988 NSFG, but could be appropriate for more recent cohorts of women for whom first marriage has been increasingly delayed.

Because the more usual indirect effect corresponds to the estimated parameter for T_1 when entered as a right-hand-side covariate in the T_2 equation, the value of this effect, under a proportional hazard specification, does not vary with time. This is reflected in the next three rows of results, which evaluate the cumulative relative risk at 90 months after onset of sexual activity for women who resided in an intact family at age 14, with the usual indirect effect identical when evaluating the cumulative relative risk at 60 and 90 months after onset. By contrast, the indirect effect of exposure declines with greater durations of exposure as initially large differences in the cumulative relative risk decrease at later exposures. These decreases are also reflected in the values for the total indirect effects, which are smaller in rows 4–6 than in rows 1–3.

The first 6 rows of Table 3 employ a single age t at which to base comparisons, using predicted medians for T_1 to generate variation in exposure to the risk of T_2 , a type of comparison depicted graphically in Figure 2. Rows 7–12 of Table 3 report indirect effects of family structure when equalizing the duration of exposure between the baseline and comparison groups (see Figure 3). These comparisons can be motivated substantively as supposing that women, on average, exit the risk of a premarital first birth not at particular ages but rather at particular durations following the sexual onset, say, if the onset of sexual activity marked the start of women’s marital search.

Equalizing the duration of exposure for the baseline and comparison groups virtually eliminates the indirect effect of exposure, with estimates of this effect close to zero.⁶ In these results, the overall indirect effect of residing in a nonintact family at age 14 is considerable, raising the relative risk of a premarital first birth by at least 25%. However, the

indirect effect of family structure due to increased exposure to risk, in which women from nonintact families are observed to have earlier entries into sexual activity (and thus longer exposure to the risk of a premarital birth), can be attributed entirely to differences in their duration at risk, and not differences in the ages at which they are at risk, since equalizing durations of exposure in Table 3 reduces these indirect effects to values close to zero. That the overall indirect effect of family structure remains large is due to the coefficient for age at first intercourse when entered as a usual right-hand-side covariate, which will reflect differences between women with different ages at sexual onset not controlled for in the other covariates in the model.

It is important to note that marriage is an implicit competing risk in our model. Thus, the results in Table 3 as well as subsequent tables should be interpreted in terms of the usual competing risk counterfactual in which we ask how prevalence would be affected by exposure to risk were women to remain unmarried.

Table 4 reports estimates for absolute and relative T_2 prevalence given in (19) and (20), respectively, paralleling the results in Table 3. Note, however, that these estimates report the consequence of both direct and indirect effects of a shift from x_1 to $x_1 + \Delta$ on T_2 prevalence.

In the first row, we compare women who resided in an intact family and a mother-only family at age 14, and trace the implications of differences in age at onset of sexual activity for the probability of a premarital first birth in 60 months following onset of sexual activity. The column labeled Pr_A shows that 7.6 percent of women from an intact family are estimated to have a premarital first birth, while the column labeled Pr_B shows that the corresponding estimate for women from a mother-only family is 17.5 percent. The absolute difference (9.9 percent) and ratio (2.30) of these two estimates are reported in the subsequent two columns.

The next two rows give estimates for the probability of a premarital first birth during the first 60 months following onset of sexual activity for women who resided at age 14 in step and other types of families. Prevalence is nearly identical for women who at age 14 resided in a step family or in other types of families, in contrast to the higher prevalence for step families relative to other families in Table 3. This result—that prevalence is more similar in Table 4 for these two groups than in Table 3—occurs for two related reasons. First, the estimates in Table 3 referred only to the indirect effect on T_2 of a shift from x_1 to $x_1 + \Delta$, while those in Table 4 combine both indirect *and* direct effects of a shift from x_1 to $x_1 + \Delta$. Second, the relative magnitude of coefficients for family structure differs for the T_1 and T_2 equations (compare columns 1 and 5 of Table 5); note, in particular, that the coefficient for having resided in a step family is larger than that for having resided in an other type of family in the T_1 equation, with this relationship reversing for the T_2 equation. As a result, the ordering of coefficients in the T_1 equation in Table 1 is directly reflected in the ordering of cumulative relative risks in Table 3, but attenuated for prevalence in Table 4.

The next three rows present estimates of absolute and relative prevalence of a premarital first birth for a 90 month period following onset of sexual activity. This longer period generates larger estimates of absolute prevalence, corresponding to longer durations at risk of the event, but slightly smaller estimates of relative prevalence.

⁶Note that in general, equalizing the duration of exposure does not imply that the indirect effect of exposure will be zero. Returning again to the thought experiment in which we compare groups *A* and *B*, it can be shown that the indirect effect of exposure will be identically zero only if the integrated hazards for groups *A* and *B* are equal, i.e., if the shaded areas are identical in the hypothetical example in Figure 3.

The final three rows of Table 4 present estimates that equalize the duration of exposure between the baseline and comparison groups. These estimates exhibit slightly smaller differences because group *B*'s exposure to risk is now more similar to group *A*'s exposure to risk, with the difference in absolute prevalence decreasing, for example, from 9.9 to 7.4 percent and that for relative prevalence from 2.30 to 1.98 when comparing rows 1 and 7.

Taken together, Tables 3 and 4 show how the cumulative relative risk and probability of a premarital first birth is affected by direct and indirect effects of family structure.⁷ Table 3 shows that the indirect effects of family structure on cumulative relative risk, while somewhat smaller than the direct effects, are still substantial, with important indirect contributions both from exposure and from the usual indirect effect of family structure as a right-hand-side covariate. Table 4 shows that the combination of indirect and direct effects increases prevalence substantially—by about 10 percent in absolute terms and by about a factor of two in relative terms.

We now turn to decomposing the direct and indirect effects, not in terms of cumulative relative risk, but in terms of the absolute and relative differences in the probability of a premarital birth. Table 5 presents a decomposition of the direct and indirect effects on absolute differences in the prevalence of a premarital first birth, using the Das Gupta decomposition technique described previously. The first row of Table 5 decomposes the absolute differences between women who resided in a mother-only family at age 14 and an intact family at age 14, evaluated 60 months after the predicted median age at first sex for women in an intact family at age 14. Under this decomposition, the direct effect of living in a mother-only family at age 14 add 5.2% to the prevalence of premarital first births. By comparison, the two indirect effects of living in a mother-only family via earlier age at first sex add 2.0% due to exposure and 2.7% due to the usual indirect effect to the prevalence of premarital first births.

The remaining rows in Table 5 mirror earlier results in Table 3, with the indirect effects generally smaller than the direct effects. In Table 5, however, we have decomposed the effects, not in terms of cumulative relative risk, but in terms of the increased percent of women who would have a premarital birth. The direct effects of the various types of nonintact family structure increase the prevalence of a premarital birth by about 5 to 10 percent relative to prevalence for those in intact families, while the pattern of results for the two indirect effects mirrors the pattern in Table 3, with the indirect of exposure declining to nearly zero values when equalizing exposure.

Table 6 presents a set of Das Gupta-style decompositions comparable to those in Table 5, but with the decompositions giving direct and indirect effects for the relative prevalence of a premarital first birth. The direct effects of residing in a nonintact family at age 14 increase the relative prevalence of a premarital birth by a factor of roughly 1.5 to 1.8. Again, the separate indirect effects are smaller than the direct effects with the indirect effect of exposure close one when equalizing exposure.

⁷In interpreting the results in both Tables 3 and 4, it is important to emphasize that we have employed estimates from a competing risk hazard model, in which women are censored if they marry prior to giving birth. While substantively sensible, a consequence is that both Tables 3 and 4 are best interpreted as speaking to a particular counterfactual—that is, the consequences of exposure on prevalence or cumulative relative risk if a woman were to remain unmarried (and hence at risk of a premarital first birth) during the entire 60 or 90 month period following onset of sexual activity. Note that this counterfactual is in contrast to the behavior of any actual sample of women, in which some percentage of women will exit the risk of a premarital birth during the 60 or 90 month period following onset of sexual activity by virtue of marriage prior to a birth, while other women will remain unmarried and at risk of a premarital birth for a much longer period than 60 or 90 months. Put another way, our empirical example assumes that endogeneities are absent from the censoring mechanism (i.e., that no endogeneity exists between marriage and a premarital first birth), an implausible assumption.

Overall, these results provide potential insights into premarital first birth risks that cannot be obtained from conventional analyses that ignore the timing of sexual onset. For example, recall that our model of premarital first birth risks conditional on first sexual intercourse allows us to specify two dimensions of time—age and duration since onset. Our empirical results suggest one way in which such a specification may be potentially useful by letting one ask whether the indirect effect of exposure can be attributed to differences in the durations at risk, in the ages at risk, both, or neither. In our empirical results, women from nonintact families are observed to have earlier entries into sexual activity and thus longer exposure to the risk of a premarital birth. In addition, our results show that equalizing durations exposed to risk reduces the indirect effect of exposure to values close to zero. Thus, we find only small differences in the probability of a premarital first birth over the five-year period following initiation of sexual activity between women who initiated sexual activity at different ages—for example, at age 17 or 19. These findings thus suggest that the indirect effect of family structure on exposure can be attributed almost entirely to differences between these groups in their durations of exposure, and not differences in the ages at which they exposed to risk. However, it is also important when weighing these observations to note that our estimates of the indirect effect of exposure are small relative to our estimates of direct effects and the more usual indirect effect.

This pattern of results—that direct effects are larger, sometimes substantially so, than the indirect effects of covariates, and that the indirect effect of exposure consistently smaller than the more usual indirect effect—is perhaps the main finding that emerges from our empirical application. Taken at face value, these results might appear contrary to the view that the longer period of abstinence in intact families is responsible for the lower numbers of premarital first births for women who grew up in such families. Likewise, these results might appear to bolster the view that premarital birth risks are most strongly affected by factors that operate in the period following sexual onset. But because our results are based on observational data, our findings should not be interpreted as providing causal evidence but rather as providing regression-adjusted estimates of various quantities. Furthermore, a proponent of the abstinence and family structure hypothesis might object that our estimates are subject to omitted variable bias.

Nevertheless, the structure of our model together with this pattern of findings implies a narrower range of relevant unobservables than is usually the case. Consider, for example, our results in Table 6 for the direct and indirect effects of having resided in a single-mother family. The estimated direct and indirect effects in Table 6 are all positive; hence, consistent with previous findings, we find that having resided in a single mother family raises the probability of a premarital first birth, with our estimated coefficients suggesting a direct effect of 56%, an indirect effect of exposure of 18%, and an indirect covariate effect of 26%. However, our estimated direct effect could well be subject to omitted variable bias.

For concreteness, consider the example of wealth as a potential candidate for an omitted w correlated with x whose omission may bias inferences about x . But because typical expectations concerning wealth would predict effects of the same sign as those for family structure, controlling for wealth would typically be expected to yield smaller estimates of both the direct and indirect effects of family structure. If so, the inclusion of an unobserved variable such as wealth would not alter the finding that the direct effect of x is larger than its corresponding indirect effect on exposure.

This example nevertheless suggests variants on standard expectations in which the inclusion of w could reverse the empirical finding that direct effects dominate indirect effects. For example, consider a w with the property that omitting w yields large biases in the estimated direct effect of x but no bias in the estimated indirect effect of x on exposure. Under this

possibility, the “true” direct effect of having resided in a single mother family could then be smaller than true indirect effect of exposure (if, for example, controlling for w were to reduce this estimate from 56% to something smaller than 18%). Note, however, that this possibility also leaves the estimate for the indirect effect of x on exposure unchanged.

It is also possible to posit conditions under which omitting w will *downwardly* bias estimates of the indirect effect of exposure. As noted above, standard expectations about w include the assumption that an omitted w will have similar signed effects as x . If this does not hold—if wealth in our example were to have a positive effect on T_1 but a negative effect on T_2 —then omission of such a w could yield a too small estimate of the indirect effect of x on exposure. But as the wealth example also illustrates, it is more difficult to posit an unobservable w correlated with x in which w has the same signed effect as x for one transition but an effect that is opposite in sign from x for the other transition.

Overall, these examples illustrate how unobserved heterogeneity can yield biased estimates of the indirect effect of exposure, but they also suggest a far narrower range of potential unobservables w than is usually the case. It is important to emphasize, however, that these observations pertain to a specific set of hypotheses—those positing a large role for T_1 exposure—and to situations in which (as is often the case) indirect effects are substantially smaller than direct effects. Note, in particular, that nothing in our framework or estimation procedures provides any novel insight concerning unobservables for hypotheses about the direct effects of covariates.

DISCUSSION

This paper has outlined a three-variable recursive proportional hazard model analogous to the classical three-variable recursive model for metric outcomes. As in the classical recursive model for metric outcomes, we posit a vector $\mathbf{x}$ of exogenous covariates, but depart from the classical model by considering two event processes T_1 and T_2 governed by a common time dimension t , in which the occurrence of T_1 is assumed to determine entry into risk of T_2 . An issue that arises in the recursive hazard model but not in the recursive linear regression model concerns the effect of exposure—variation in T_1 will generate differential exposure to risk, and hence differential prevalence in the event T_2 . We show that under proportionality, one can decompose the cumulative relative risk of T_2 conditional on T_1 into two multiplicative components, one of which reflects an effect of exposure and the other of which reflects an effect analogous to the traditional indirect effect in linear regression. Although this multiplicative decomposition does not obtain when comparing relative or absolute prevalence, we provide derivations for direct and indirect effects of $\mathbf{x}$ on T_2 for absolute and relative measures of prevalence, as well as for the cumulative relative risk.

Because our method requires that one be able to evaluate the integral of the baseline hazard function, it is most easily applied to parametric models of the baseline hazard function. Our examples use a splined piecewise Gompertz specification for the baseline hazard, which is a highly flexible model for the baseline hazard, but other parametric choices can be utilized as well. Note, however, that use of a Cox proportional hazard model would present difficulties because it does not provide direct estimates of the baseline hazard function and, in particular, estimates suited to inverting the integral of the baseline hazard function.

We illustrate these methods using data on age at first intercourse and age at a premarital first birth, estimating the direct and indirect effects of family structure on the age-specific risk of a premarital birth. Our empirical analyses suggest that roughly one-half of the effect of nonintact family structure on the risk of a premarital first birth is due to the conventional direct effect of family structure, in which the risk of a premarital first birth among sexually

active women is higher for women from nonintact families. However, we also found that between one-fourth and one-half of the effect of nonintact family structure on the risk of a premarital first birth is due to an indirect effect of family structure, in which family structure influences the risk of a premarital first birth indirectly through its effect on age at first sexual intercourse. The largest component of this indirect effect is that women from nonintact families have an earlier onset of sexual activity, with a woman's age at first sex, entered as a right-hand-side covariate for women's risk of a premarital first birth, associated with higher risks of a premarital first birth. An additional indirect effect, which arises in a hazard context, is that an earlier onset of sexual activity will increase a woman's duration of exposure to the risk of a premarital first birth. In all our examples, this component was the smallest of the effects of family structure on premarital first birth risks.

Thus, a consistent pattern that emerges from our empirical decompositions is that the direct effects of covariates are, in general, larger in magnitude than their corresponding indirect effects. Although these findings will for many social scientists be unsurprising—direct effects very often dominate indirect effects—because our data are observational, it is not possible to regard these findings as causal, but rather as providing estimates of various regression-adjusted quantities. However, we also show that unobservables must have highly specific characteristics if indirect effects are to dominate direct effects, a result that follows from the structure of our formal model. This implies a far smaller class of relevant unobserved confounds than is usually the case for hypotheses positing large indirect effects of exposure.

More generally, our empirical example of how one can decompose the direct and indirect effects of family structure on nonmarital childbearing illustrates the potential of this decomposition technique for modeling many other event processes. Whenever a hazard model outcome can be viewed in terms of two linked event processes, one gains the ability to distribute the effects of a covariate across the two event processes. By distinguishing between direct and indirect effects of the covariate, and by further distinguishing between a conventional indirect effect *and* an indirect effect related to duration of exposure, researchers can gain additional insight into the circumstances and dynamics under which an explanatory variable is associated with a particular outcome.

Acknowledgments

We gratefully acknowledge funding from NICHD (HD 29550) and helpful comments from the editor and anonymous reviewers. Earlier versions of this paper also benefitted from comments by participants at the 2008 Low Income Workshop, Institute for Poverty, University of Wisconsin-Madison.

APPENDIX 1

Obtaining the indirect effect of $\mathbf{x}$ on T_2 prevalence requires inverting the integral of $q_1(t)$. The examples in this paper employ a piecewise splined Gompertz specification for the baseline hazard $q_1(t)$ of T_1 . Under proportionality, we have

$$r_1(t) = q_1(t) \exp(\mathbf{a}\mathbf{x}). \quad (\text{A1})$$

Consider partitioning the time interval (τ_0, ∞) into K prespecified intervals $(\tau_0, \tau_1]$, $(\tau_1, \tau_2]$, ..., $(\tau_{K-1}, \infty]$; then a piecewise splined Gompertz specification for $q_1(t)$ can be written as:

$$q_1(t) = \begin{cases} \exp(\beta_1 + \gamma_1 t) & t \in (\tau_0, \tau_1]; \\ \exp(\beta_2 + \gamma_2 t) & t \in (\tau_1, \tau_2]; \\ \dots & \\ \exp(\beta_K + \gamma_K t) & t \in (\tau_{K-1}, \infty], \end{cases} \quad (\text{A2})$$

for $\gamma_k \neq 0$ and $k = 1, \dots, K$. Under (A1) and (A2), $H(t)$ is given by:

$$H_1(t) = \int_{\tau_0}^t r_1(s) ds = \exp(\mathbf{a}\mathbf{x}) \int_{\tau_0}^t q_1(s) ds = \exp(\mathbf{a}\mathbf{x}) Q_1(t). \quad (\text{A3})$$

Integrating $Q_1(t)$ yields:

$$Q_1(t) = \begin{cases} e^{\beta_1 + (\gamma_1 t - e^{\gamma_1 \tau_1}) / \gamma_1} & t \in (\tau_0, \tau_1]; \\ Q_1(\tau_1) + e^{\beta_2 + (\gamma_2 t - e^{\gamma_2 \tau_2}) / \gamma_2} & t \in (\tau_1, \tau_2]; \\ \dots & \\ Q_1(\tau_{K-1}) + e^{\beta_K + (\gamma_K t - e^{\gamma_K \tau_K}) / \gamma_K} & t \in (\tau_{K-1}, \infty]. \end{cases} \quad (\text{A4})$$

As noted in the text, our goal is to determine the p percentile of the T_1 distribution; this corresponds to inverting the integral of $q_1(t)$. Set $Q_1(t) = x$ and define the inverse function implicitly through $Q_1^{-1}(x) = t$. Suppose the desired percentile lies in the k th interval $(\tau_{k-1}, \tau_k]$; then from (A4)

$$\begin{aligned} x &= Q_1(\tau_{k-1}) + [\exp(\beta_k + \gamma_k t) - \exp(\beta_k + \gamma_k \tau_{k-1})] / \gamma_k \\ \exp(\beta_k + \gamma_k t) &= \exp(\beta_k + \gamma_k \tau_{k-1}) + \gamma_k [x - Q_1(\tau_{k-1})] \\ t &= [\log(\exp(\beta_k + \gamma_k \tau_{k-1}) + \gamma_k [x - Q_1(\tau_{k-1})]) - \beta_k] / \gamma_k \end{aligned} \quad (\text{A5})$$

Hence for $t \in (\tau_{k-1}, \tau_k]$,

$$Q_1^{-1}(x) = [\log(\exp(\beta_k + \gamma_k \tau_{k-1}) + \gamma_k [x - Q_1(\tau_{k-1})]) - \beta_k] / \gamma_k. \quad (\text{A6})$$

Minor complications arise when the distribution of T_1 is defective—that is, when some individuals will not experience the event T_1 even when $t \rightarrow \infty$. For the piecewise Gompertz specification, this is determined by parameters in the last open interval, $(\tau_{K-1}, \infty]$. In this interval, the T_1 distribution will be defective if

$$\gamma_K e^{-\beta_K} [x - Q_1(\tau_{K-1})] < e^{\gamma_K \tau_{K-1}}. \quad (\text{A7})$$

Inspecting (A7) shows that $\gamma_K < 0$ is a necessary but not sufficient condition for the distribution of T_1 to be defective.

APPENDIX 2

In this appendix, we sketch details when there are time-varying covariates. For the transition to T_1 , we seek, as before, to make comparisons between groups A and B using the p th percentile of the distribution of T_1 , t_{1p} . To fix ideas, consider the case of a single time-varying covariate, $x_1(t)$. Then a natural means of comparison for groups A and B (see, e.g., self-citation 5) is to compare two (often hypothetical) trajectories for x_1 , with the remaining

covariates set to their means, with $\tilde{x}_{1A}(t)$ and $\tilde{x}_{1B}(t)$ denoting these two hypothetical trajectories for $x_1(t)$. This then generalizes to $\tilde{\mathbf{x}}_{1A}(t)$ and $\tilde{\mathbf{x}}_{1B}(t)$, with the vector $\mathbf{x}(t)$ comprised of a mix of non-time-varying and time-varying covariates. The expressions in (7) and (8) then need modification as:

$$t_{1p} = Q_1^{-1} \left[-\log(\pi) / \exp(\mathbf{a} \tilde{\mathbf{x}}(t)) \right], \quad (7.A2)$$

and

$$\begin{aligned} \Delta t_{1p} &= t_{1p}^B - t_{1p}^A \\ &= Q_1^{-1} \left\{ -\log(\pi) / \exp[\mathbf{a} \tilde{\mathbf{x}}_B(t)] \right\} - Q_1^{-1} \left\{ -\log(\pi) / \exp[\mathbf{a} \tilde{\mathbf{x}}_A(t)] \right\}, \end{aligned} \quad (8.A2)$$

where, as before, t_{1p}^A and t_{1p}^B denote the p th percentiles of T_1 in groups A and B , respectively, and where the expression for $Q_1^{-1}(t)$ now requires integration over $\tilde{\mathbf{x}}(t)$.

Other expressions follow similarly.

Appendix Table 1

Comparison of estimated covariate effects from splined piecewise Gompertz and Cox models for: (a) the unconditional transition to a premarital first birth, (b) the transition to first sexual intercourse, and (c) the transition to a premarital first birth conditional on onset of sexual activity. White women, National Survey of Family Growth, 1988.

	Unconditional transition to a premarital first birth (no adjustment for left truncation)				Transition to first sexual intercourse		Transition to a premarital first birth given sexual onset	
	Gmp	Cox	Gmp	Cox	Gmp	Cox	Gmp	Cox
mother-only family	.96*** (.16)	.96*** (.16)	.40* (.17)	.40* (.17)	.51*** (.06)	.50*** (.06)	.47** (.16)	.48** (.16)
step family	1.06*** (.20)	1.05*** (.20)	.66** (.20)	.66** (.20)	.69*** (.07)	.67*** (.07)	.58** (.20)	.58** (.20)
other family	1.04*** (.20)	1.03*** (.20)	.75*** (.20)	.74*** (.20)	.58*** (.07)	.58*** (.07)	.64** (.20)	.62** (.20)
catholic	-.23 (.14)	-.23 (.14)	-.15 (.14)	-.16 (.14)	-.01 (.04)	-.01 (.04)	-.22 (.14)	-.22 (.14)
mother's education	-.08** (.02)	-.07** (.02)	-.09*** (.02)	-.09*** (.02)	.01 (.01)	.01 (.01)	-.10*** (.02)	-.10*** (.02)
mother's age at first birth	-.10*** (.02)	-.10*** (.02)	-.06** (.02)	-.06** (.02)	-.05*** (.00)	-.05*** (.00)	-.06** (.02)	-.06** (.02)
age at first intercourse			-.026*** (.002)	-.026*** (.002)			-.018*** (.004)	-.016*** (.004)
duration 32–59 months							-.40* (.20)	-.28 (.21)
duration 60+ months							-1.01* (.40)	-.77 (.44)

* $p < .05$

** $p < .005$

*** $p < .0005$ (two-tailed tests)

REFERENCES

- Aalen, Odd O. Nonparametric Inference for a Family of Counting Processes. *Annals of Statistics*. 1978; 6(4):701–26.
- Breslow NE. Covariance Analysis of Censored Survival Data. *Biometrics*. 1974; 30(1):89–99. [PubMed: 4813387]
- Cox DR. Regression Models and Life Tables” (with discussion). *Journal of the Royal Statistical Society*. 1972; B34(2):187–220.
- Das Gupta, Prithwis. Current Population Reports Special Studies P-23, No. 186. Bureau of the Census; Washington, DC: 1993. Standardization and Decomposition of Rates: A User’s Manual.
- Duncan, Otis Dudley. Introduction to Structural Equation Models. Academic Press; New York: 1975.
- Featherman, David L.; Hauser, Robert M. Opportunity and Change. Academic Press; New York: 1978.
- Hamilton, Brady E.; Martin, Joyce A.; Ventura, Stephanie J. National Vital Statistics Reports. Vol. Volume 56. National Center for Health Statistics; Rockville, MD: 2007. Births: Preliminary Data for 2006.
- Kalbfleisch, John D.; Prentice, Ross L. The Statistical Analysis of Failure Time Data. Wiley; New York: 1980.
- Kiernan, Kathleen E.; Hobcraft, John. Parental Divorce during Childhood: Age at First Intercourse, Partnership, and Parenthood. *Population Studies*. 1997; 51(1):41–55.
- Kitagawa, Evelyn M. Components of a Difference Between Two Rates. *Journal of the American Statistical Association*. 1955; 50(272):1168–94.
- Lillard, Lee. Simultaneous Equations for Hazards: Marriage Duration and Fertility Timing. *Journal of Econometrics*. 1993; 56(1/2):189–217. [PubMed: 12318328]
- Smith, Herbert L.; Morgan, S. Philip; Koropecj-Cox, Tanya. A Decomposition of Trends in the Nonmarital Fertility Ratios of Blacks and Whites in the United States, 1960 to 1992. *Demography*. 1996; 33(2):141–52. [PubMed: 8827161]
- Stolzenberg, Ross M. The Measurement And Decomposition Of Causal Effects In Non-linear And Nonadditive Models. In: Schuessler, Karl, editor. *Sociological Methodology 1980*. Jossey-Bass; San Francisco: 1979. p. 459-88.
- Winship, Christopher; Mare, Robert D. Regression Models with Ordinal Variables. *American Sociological Review*. 1984; 49(4):512–25.

Figure 1.

Smoothed nonparametric estimates of: (a) age dependence in the logarithm of the hazard for the transition to first sexual intercourse, (b) age dependence in the transition to a premarital first birth, and (c) duration dependence in the transition from sexual onset to a premarital first birth. White women, National Survey of Family Growth 1988.

Figure 2. Hypothetical indirect effect of Δ a change in t_{1A} when evaluated at a fixed age t when integrating the hazard rate for T_2 for the baseline and comparison groups.

Figure 3.
Hypothetical indirect effect of a Δ change in t_{1A} when evaluated at the same duration since onset when integrating the hazard rate for T_2 for the baseline and comparison groups.

Table 1

Effects of family structure on the age-specific rate of first sexual intercourse, a premarital first birth ignoring exposure since onset of first sexual intercourse, and a premarital first birth correcting for exposure since onset of first sexual intercourse. White women, National Survey of Family Growth 1988.

	First sex		Premarital first birth			
	no adjustment for left truncation		adjusting for left truncation			
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Family Structure</i>						
mother-only family	.51*** (.06)	.96*** (.16)	.40* (.17)	.58*** (.16)	.48** (.16)	.47** (.16)
step family	.69*** (.07)	1.06*** (.20)	.66** (.20)	.63** (.20)	.58** (.20)	.58** (.20)
other family	.58*** (.07)	1.04*** (.20)	.75*** (.20)	.66** (.20)	.62** (.20)	.64** (.20)
<i>First Sexual Intercourse</i>						
age in months			-.026*** (.002)		-.010*** (.002)	-.018*** (.004)
<i>Duration since sexual onset</i>						
32 to 71 months						-.40* (.20)
72+ months						-1.01* (.40)

Note: All models include an age-specific baseline function and include controls for background variables. Column (1) presents coefficients for the age-specific risk of sexual onset, while columns (2)–(6) present coefficients for the age-specific risk of a premarital first birth. Columns (2) and (3) do not model exposure since onset of first sexual intercourse; columns (4)–(6) model such differences by left truncating premarital first birth risks by a woman's age at first sexual intercourse. Columns (3), (5) and (6) include a control for age at first sexual intercourse as a right-hand-side covariate. Column (6) gives estimates for the age- and duration-specific risk of a premarital first birth conditional on entry into sexual activity. See text for additional details.

*
 $p < .05$

**
 $p < .005$

 $p < .0005$ (two-tailed tests)

Table 2

Predicted median age (in months) at first sexual intercourse for women residing in intact, mother-only, step, and other families at age 14 and predicted difference relative to women residing in an intact family at age 14. White women, National Survey of Family Growth 1988.

	age (in months)	diff
<i>Family structure</i>		
intact	226.5	0.0
mother-only	213.0	13.4
step	209.1	17.4
other	211.3	15.2

Decomposition of direct and indirect effects (in percentage change) for the cumulative relative risk of a premarital first birth for white women under alternative simulations for the median age at first sexual intercourse by family structure at age 14.

Table 3

Simulation	Intact family		Comparison		Percent change in cumulative relative risk			
	t_{1A}	t	t_{1B}	t	direct effect	indirect exposure	usual indirect effect	total indirect effect
Comparison group								
1. mother-only	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + 60$	60.4	19.0	27.5	51.8
2. step	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + 60$	79.2	22.8	35.8	66.8
3. other	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + 60$	89.1	20.9	31.4	59.0
4. mother-only	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + 90$	60.4	14.4	27.5	45.9
5. step	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + 90$	79.2	17.3	35.8	59.3
6. other	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + 90$	89.1	15.9	31.4	52.3
7. mother-only	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 60$	60.4	0.6	27.5	28.3
8. step	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 60$	79.2	-0.2	35.8	35.5
9. other	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 60$	89.1	0.3	31.4	31.8
10. mother-only	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 90$	60.4	-0.7	27.5	26.6
11. step	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 90$	79.2	-1.6	35.8	33.6
12. other	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 90$	89.1	-1.1	31.4	30.0

Note: Calculations based on estimates from column 6 of Table 1 and on t_{1A} = predicted median age at first sexual intercourse for women who resided in an intact family at age 14 as reported in Table 2. See text for additional details.

Absolute and relative differences for the probability of a premarital first birth for white women under alternative simulations for the median age at first sexual intercourse by family structure at age 14.

Table 4

Simulation	Intact family		Comparison		Absolute and relative differences in the probability of a premarital first birth			
	t_{1A}	t	t_{1B}	t	Pr_A	Pr_B	$Pr_B - Pr_A$	Pr_B / Pr_A
Comparison group								
1. mother-only	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + 60$	7.6	17.5	9.9	2.30
2. step	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + 60$	7.6	21.0	13.4	2.77
3. other	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + 60$	7.6	21.1	13.5	2.78
4. mother-only	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + 90$	9.9	21.6	11.7	2.19
5. step	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + 90$	9.9	25.7	15.8	2.60
6. other	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + 90$	9.9	25.9	16.0	2.62
7. mother-only	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 60$	7.6	15.0	7.4	1.98
8. step	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 60$	7.6	17.4	9.8	2.30
9. other	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 60$	7.6	17.8	10.3	2.35
10. mother-only	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 90$	9.9	19.1	9.2	1.93
11. step	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 90$	9.9	22.1	12.2	2.23
12. other	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 90$	9.9	22.6	12.7	2.28

Note: Calculations based on observed sample means, estimates in column 6 of Table 1, and t_{1A} = predicted median age at first sexual intercourse for women who resided in an intact family at age 14 as reported in Table 2. See text for additional details.

Decomposition of direct and indirect effects for the absolute prevalence of a premarital first birth for white women under alternative simulations for the median age at first sexual intercourse by family structure at age 14.

Table 5

Simulation	Intact family		Comparison		Absolute change in prevalence			
	t_{1A}	t	t_{1B}	t	direct effect	indirect effect of exposure	usual indirect effect	total effect
Comparison group								
1. mother-only	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + 60$	5.2	2.0	2.7	9.9
2. step	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + 60$	7.1	2.5	3.8	13.4
3. other	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + 60$	7.8	2.4	3.4	13.5
4. mother-only	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + 90$	6.5	1.9	3.4	11.7
5. step	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + 90$	8.7	2.4	4.6	15.8
6. other	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + 90$	9.6	2.3	4.2	16.0
7. mother-only	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 60$	4.8	0.1	2.5	7.4
8. step	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 60$	6.4	0.0	3.4	9.8
9. other	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 60$	7.1	0.0	3.1	10.3
10. mother-only	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 90$	6.1	-0.1	3.2	9.2
11. step	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 90$	8.1	-0.2	4.3	12.2
12. other	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 90$	9.0	-0.2	3.9	12.7

Note: Calculations based on estimates in column 6 of Table 1 and on t_{1A} = predicted median age at first sexual intercourse for women who resided in an intact family at age 14 as reported in Table 2. See text for additional details.

Table 6

Decomposition of direct and indirect effects for the relative prevalence of a premarital first birth for white women under alternative simulations for the median age at first sexual intercourse by family structure at age 14.

Simulation	Intact family		Comparison		Relative change in prevalence			
	t_{1A}	t	t_{1B}	t	direct effect	indirect effect of exposure	usual indirect effect	total effect
Comparison group								
1. mother-only	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + 60$	1.56	1.18	1.26	2.30
2. step	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + 60$	1.72	1.21	1.33	2.77
3. other	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + 60$	1.81	1.19	1.29	2.78
4. mother-only	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + 90$	1.55	1.13	1.25	2.19
5. step	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + 90$	1.70	1.16	1.32	2.60
6. other	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + 90$	1.79	1.14	1.28	2.62
7. mother-only	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 60$	1.56	1.01	1.26	1.98
8. step	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 60$	1.73	1.00	1.33	2.30
9. other	t_{1A}	$t_{1A} + 60$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 60$	1.82	1.00	1.29	2.35
10. mother-only	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 90$	1.55	0.99	1.25	1.93
11. step	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 90$	1.71	0.99	1.32	2.23
12. other	t_{1A}	$t_{1A} + 90$	$t_{1A} + \Delta$	$t_{1A} + \Delta + 90$	1.80	0.99	1.28	2.28

Note: Calculations based on estimates in column (5) of Table 1 and on t_{1A} = predicted median age at first sexual intercourse for women who resided in an intact family at age 14 as reported in Table 2. See text for additional details.