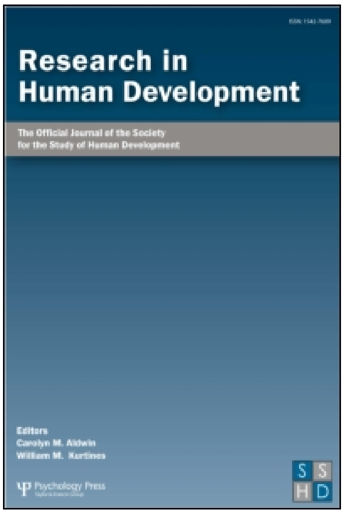

Research in Human Development
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t775653695

Understanding Long-Term Diffusion Dynamics in the Prevalence of Adolescent Sexual Initiation: A First Investigation Using Agent-Based Modeling

Mark G. Orr[a]; Clare Rosenfeld Evans[a]
[a] Columbia University,

Online publication date: 23 February 2011

To cite this Article Orr, Mark G. and Evans, Clare Rosenfeld(2011) 'Understanding Long-Term Diffusion Dynamics in the Prevalence of Adolescent Sexual Initiation: A First Investigation Using Agent-Based Modeling', Research in Human Development, 8: 1, 48 — 66

To link to this Article: DOI: 10.1080/15427609.2011.549702

URL: http://dx.doi.org/10.1080/15427609.2011.549702

Research in Human Development, *8*(1), 48–66, 2011

ISSN: 1542-7609 print/1542-7617 online
DOI: 10.1080/15427609.2011.549702

Psychology Press
Taylor & Francis Group

Understanding Long-Term Diffusion Dynamics in the Prevalence of Adolescent Sexual Initiation: A First Investigation Using Agent-Based Modeling

Mark G. Orr and Clare Rosenfeld Evans
Columbia University

Time-trends in the prevalence of adolescent sexual initiation exhibit periods of increase, decrease and equilibrium. We attempted to explain, in mechanistic terms, how such dynamics arise by developing an Agent-Based Model. The model assumes that sexual initiation diffuses socially both within and across cohorts. The model behavior matched, qualitatively, the empirical time-trends. The impact of two intervention strategies suggested that the age at which an intervention is implemented effected system behavior as did the choice of which specific subpopulation was targeted. Suggestions for how computational models might be used to explore research questions in developmental science were discussed.

Social diffusion—the idea that behaviors, attitudes, beliefs, information, and innovations can spread socially—is a fundamental explanatory mechanism that is foundational to several of the social and behavioral sciences (e.g., diffusion of innovations [Rogers, 2003]), social psychology (e.g., social cognitive theory [Bandura, 1997]), sociology (e.g., social network theory, [Newman, 2010]) and social epidemiology (e.g., theory of fundamental causes of disease, [Link, 2008]). Furthermore, social diffusion is well supported empirically (see Smith & Christakis, 2008, for a review); for example, it has been implicated in smoking cessation (Christakis & Fowler, 2008), emotions (Hill, Rand,

Address correspondence to Mark G. Orr, Department of Population and Family Health, Mailman School of Public Health, Columbia University, 60 Haven Avenue B-2, New York, NY 10032. E-mail: mo2259@columbia.edu

Nowak, & Christakis, 2010a), mental health (Ueno, 2005), drug use (Galea, Hall, & Kaplan, 2009), suicidal ideation (Bearman & Moody, 2004), and altruistic cooperation (Fowler & Christakis, 2010). In practice, public health interventions that capitalize on diffusion principles have demonstrated success (Valente, 2010).

Progress has been made toward understanding several important aspects of social diffusion, including the probability of an individual being influenced by each of his or her social contacts (Hill, et al., 2010a; Hill, Rand, Nowak, & Christakis, 2010b), the introduction of individual-level thresholds as the basis for diffusion dynamics (Granovetter, 1978), how the rate of social diffusion is dependent on network structures and patterns of thresholds (Newman, 2010; Valente, 1995; Watts, 2002), the diffusion dynamics between dyadic relations (Shoda, LeeTiernan, & Mischel, 2002), and the potential influence of social diffusion on population-level distributions of personality traits (Read et al., 2010). The bulk of this work has addressed static populations, where the members of the population under study do not change. Therefore, little has been revealed thus far about diffusion mechanisms across larger time scales whereby populations are separated in time.

Larger time scales (e.g., 10, 20, 40 years) are important to consider because the trends in prevalence of several key behaviors (smoking, obesity, consumer behavior, sexual behavior, public opinion) are suggestive of a diffusion process that occurs between populations separated in time (what we call intercohort diffusion). For example, the prevalence of adult cigarette smoking in the United States dropped every time it was measured from 1965 to 2007 (Centers for Disease Control and Prevention, 2010). This trend is not suggestive of a random change in smoking behavior across cohorts, but of some systematic influence on smoking behavior. That is to say, each successive cohort was pushed to smoke less than previous cohorts. Given that social diffusion is well recognized within single cohorts and that birth cohorts have some overlap (i.e., exist simultaneously), we hypothesize that intercohort diffusion may be one of the driving forces that creates systematic change versus random fluctuations across cohorts.

In this article, we approach this difficult problem by conceptualizing it as a very simplified system. Our approach, although a gross oversimplification of the problem, will afford the initial traction needed to generate potential insights into the phenomenon of interest. Figure 1 represents an age-graded two-cohort system, where Bin A contains a cohort that is younger than the cohort contained in Bin B. The mechanics of the system are as follows. A cohort enters the system in Bin A, ages for a fixed time period, and then migrates to Bin B where it ages for another fixed time period before finally exiting the system. When a cohort migrates to Bin B, another younger cohort enters into Bin A. Within each bin the social diffusion process is mostly governed by interactions within the bin. However,

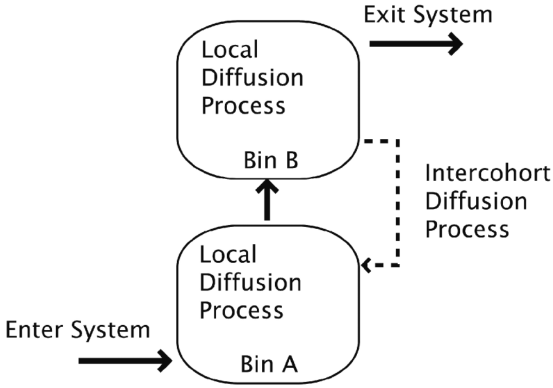

FIGURE 1 Conceptual schematic of intercohort diffusion of behavior model (IDBM) diffusion processes. This conceptual schematic of diffusion processes in the IDBM depicts three critical aspects of the model. First, the portion of the population under study at a given point in time is segmented into the older and younger cohorts. Second, diffusion largely occurs between individuals in the same cohort who are local neighbors. Third, diffusion also occurs from the older cohort to select individuals in the younger cohort who are linked to the older population of agents. The younger cohort is assumed not to have any influence on the behavior of the older cohort. Each cohort of agents enters Bin A and spends four time ticks (nonspecified units of time) there before progressing to Bin B where the cohort remains for another four time ticks before exiting the system. As the first cohort move from Bin A to Bin B, the subsequent cohort enters Bin A and follows the same process. Each cohort is only active in the system for a total of 8 ticks (4 in Bin A and 4 in Bin B) and cohorts proceed through the system sequentially.

some members in Bin A are influenced by the members in Bin B (this is the intercohort diffusion process). The measure of interest in this case would be the prevalence of the behavior in Bin B, measured once for each cohort. This provides a time series, cross-sectionally by cohort, of the prevalence of the behavior in the system.

Dynamics of Sexual Initiation Behavior in Adolescents

Our interest in intercohort diffusion stems from the long-term time trends of adolescent sexual initiation prevalence in the United States from about 1950 to the present (sexual initiation refers to the first time a person engages in coitus). These data show three empirical regularities: (1) periods of monotonic (unidirectional) increases or decreases in prevalence; (2) change points, or points in time at which the prevalence time trend shifts direction; and (3) periods of equilibria in which the prevalence is static, not increasing or decreasing. For example, among 18-year-old women in the United States, prevalence rose monotonically from 26%

to 54% during the period between the late 1950s and approximately 2000 (Finer, 2007). Data from U.S. high school students showed a monotonic decrease in sexual experience during the 1990s and early 2000s that has more recently reached a phase of little change in prevalence over time (Santelli, Orr, Lindberg, & Diaz, 2009).

These time trends have been collected mainly for the adolescent sexual and reproductive health community in an effort to understand past trends and to possibly prepare for future health needs in contraception, health education, and reproductive services. The focus of this literature, however, has not been toward developing theoretical or mechanistic explanations of these time trends (see Singh & Darroch, 1999, pp. 218–219 for a prototypical example of how mechanisms are not considered). The purpose of this article is to begin investigating the options for a mechanistic, theoretical account of the long-term trends in sexual initiation in the United States. This will not only increase our scientific understanding of the etiology of the trends, but may also allow for better preparation for future reproductive health needs.

To this end, we developed the intercohort diffusion of behaviors model (IDBM) to explore whether long-term dynamics could arise in a system that is constrained to include only those structures and mechanisms presented in Figure 1. We chose to focus on social diffusion because of its ubiquity across health-related behaviors and its specific role in sexual initiation in adolescents (Rodgers, Rowe, & Buster, 1998; Romer & Stanton, 2003). Our work is not intended to discount the importance of exogenous influences such as technological advancements in contraceptives (Coontz, 1992) or the influence of the mass media (Brown, 2002). In contrast, we intend to provide a proof of concept that long-term dynamics may arise without such exogenous influences.

COMPLEX SYSTEMS AND AGENT-BASED MODELING

Complex systems research "challenges the notion that by perfectly understanding the behavior of each component part of a system we will then understand the system as a whole" (Miller & Page, 2007, p.3). The behavior of complex systems is described well by a set of characteristic features, of which the most applicable to the IDBM are sensitivity to initial conditions, bifurcation (dramatic shifts in the observed behavior of the system), and self-organization. *Self-organization* refers to the global, system-level patterns that are driven by a process of interaction among lower-level elements in the system.

Agent-based modeling (or ABM) is a computational tool that is well suited to understanding complex systems. In ABM the interactions of individuals—called agents—are governed strictly by local interaction rules that are not determined by the global, system-level behavior or by any influence external to the system.

An alternative methodology to ABM, which is powerful and popular throughout the social sciences, is differential equation modeling (e.g., it is the backbone of system dynamics modeling). We employed ABM because key features of real world structures—including heterogeneity and network structure—would be best captured through the use of this framework (Rahmandad & Sterman, 2008). Furthermore, we felt that potential forthcoming expansions of the IDBM may be more amenable to the ABM framework (see Discussion for forthcoming expansions of the IDBM).

The behavior of the IDBM is expected to accord with what is already known about complex systems. In particular, we expect to observe global (systems level) patterns of behavior arising from self-organization (individual-level behavior), bifurcation (phase shifts), and sensitivity to initial conditions. That is, we expect to see changes in population-level prevalence that arise out of individual-level sexual behavior, and that the observed prevalence will depend on what conditions are assumed to exist at the outset.

METHOD

Overview of the IDBM

Before absorbing details of the IDBM, it will be helpful to have a skeletal understanding of how cohorts flow into and out of the model and how the diffusion process operates within and between cohorts. We provide the details in the next section for all aspects of the IDBM. In particular, the terms *social context, social influence*, and *how time is handled* will be defined explicitly. It will be helpful to consult Table 1 (a listing of parameters and their justifications in the IDBM).

Figure 1 is a schematic of the structure and dynamical flow of the IDBM. A cohort of agents (representative of a cohort of individual people) enters the system into Bin A, stays in Bin A for a specified time-period, migrates to Bin B as a cohort, stays for a specified time-period before exiting the system. When a cohort migrates to Bin B, Bin A is populated with a new cohort. This population flow mimics how grades are structured in a school.

This initial model is a deliberate oversimplification. Thus, we did not represent age, race/ethnicity, or gender as properties of cohorts or individual agents. These are known factors that show differences in the prevalence of sexual initiation in adolescents.

In the IDBM each agent can be in one of two states—sexually initiated or not (this is a one-way state system; once sexually initiated, an agent cannot switch back to noninitiated). Social diffusion in the IDBM, then, is the spread of becoming sexually initiated and is governed by the behavior of the individual

TABLE 1
Summary of Parameters and Justifications in the IDBM

Parameters of Model*	
Social Threshold Mean & Standard Deviation	The use of a normal distribution in threshold models reflects the empirical findings in the diffusion of innovations literature (Rogers, 2003). The mean and standard deviation of this distribution are two separate model parameters. Our selection of the distribution $N(0.40,0.12)$ was based in part on our sensitivity analysis during the parameter selection phase (see *Figure 3*, row D) where we tested mean values ranging from 0.38 to 0.42. The selection of sd $= 0.12$ was arbitrary. It is critical to note that a threshold model does not specify how a given agent came to have a particular threshold.
% Linked	The values explored in this paper were 12%, 14% and 16%. These values are derived from the (somewhat limited) empirical evidence available on typical inter-grade connectedness in adolescent school environments (Moody, 1999).
Gamma (γ)	For those agents in the younger cohort (*Bin A*) that are linked to the older cohort (*Bin B*), γ represents the weight placed on the influence from the older cohort (See Eq. 2). The range of γ is 0.0 to 1.0 with 0.0 representing no influence (and thus essentially the agent is unlinked) and 1.0 indicating that only the older cohort—not neighboring agents in the younger cohort—is influential.
Sexually Initiated at Model Initialization	A fixed proportion of each new cohort is assigned the status of "sexually initiated " at model initialization in order to trigger diffusion within the cohort. These "seed" agents are selected at random. The proportion used in all simulations was 8%. The choice to use 8% was necessarily arbitrary given that no "real world" age for the group was defined. The proportion of a population that would be a realistic seed depends heavily on the age of that population when observation begins.
Number of Agents per Cohort	In all simulations we used 650 agents in each cohort. This selection was arbitrary but allowed for sufficient social connectedness to observe interesting system behavior.
Area of Physical Environment	The area of the physical environment impacts the density of the groupings of agents and thus the social connectedness (i.e. the average number of neighbors for each agent). By default NetLogo creates an environment of 33×33 equal-sized square spaces and we did not vary this parameter. The 33×33 space was divided horizontally to create the two 16×33 spaced *Bins.*

(*Continued*)

TABLE 1
(*Continued*)

Summary of Heterogeneous Parameters	There are a number of ways individual agents differ from each other: (1) individual social thresholds (2) number of neighbors (3) status as sexually initiated or not (4) status as *linked* to the older cohort or not.
Model Dynamics	
Spatial Distribution & Network Structure	Physical location of agents is assigned randomly when the cohort is created. Since physical proximity is used as a proxy for social connectedness, the random assignment of location creates heterogeneity in the number of social contacts a given agent possesses as well as heterogeneity in overall network structure between cohorts. Our choice to use random spatial distribution as opposed to the more even spacing of agents in a lattice and the assignment of social connections through a probability function, as some other researchers do, was arbitrary.
Run	All runs were preset to run for 2,500 ticks. The selection of 2,500 ticks was arbitrary yet we deemed it sufficient to allow a significant number of cohorts to progress through the artificial environment. Given that all simulations had a run time of 2,500 ticks there were a total of 625 cohorts visible during this interval. The last cohort, however, only completes the four ticks in younger grade so the total is reduced to 624.

*The term parameter refers to aspects of the IDBM that are under the control of the modeler. They serve as constants when the model is running, but can be manipulated between runs of the model.

agents. The switch for each agent to the state *sexually initiated* depends on its social threshold function—the input for the social threshold function is the amount of social influence to initiate sexual activity that comes from the agent's social context. If at any time the agent's social influence exceeds the agent's social threshold, the agent changes states from not sexually initiated to sexually initiated. For most agents, social context is defined within the agent's cohort. For a handful of agents in Bin A, social context is defined within the agent's cohort and the cohort in Bin B (this drives intercohort diffusion).

Specifics of the IDBM

Social context. To capture local effects of agent-to-agent social influence, each bin was represented by a set of 528 squares arranged on a 16 by 33 Cartesian coordinate system (see area of physical environment in Table 1). A cohort consisted of 650 agents (see number of agents per cohort in Table 1), each of which was randomly assigned to reside permanently in one square when the agent's cohort entered the system (multiple agents could reside on the same square).

The social context for most agents was composed of all agents on adjacent squares (called a neighborhood) within a bin. This amounted to 8, 5, or 3 squares, depending on where the agent resided—for example, if the agent was in one of the corners, it had three adjacent squares. For a small fraction of agents in Bin A, social context included their neighborhood plus influence from Bin B (this will be explained under Agent Social Threshold Function). These agents are called linked agents. Linked status was assigned randomly when agents entered the cohort. The fraction of linked agents was determined by our sensitivity analysis (see Parameter Selection and Sensitivity Analysis below).

Dynamics, time, and measurement. Time is represented in the model by discrete, arbitrary units called ticks. The tick coordinates the social diffusion process. A tick is a signal to all agents in the system to update their state, synchronously, using the social threshold function (see section Agent Social Threshold Function below).

A cohort enters the system (into Bin A) and stays there for four ticks. Then, it migrates to Bin B and stays for four ticks before exiting the system. When a cohort moves to Bin B another cohort enters Bin A. Thus, except for the first cohort that enters the system, two cohorts always exist simultaneously in the system for 4 ticks (one cohort in Bin A; the other in Bin B).

A run consisted of putting 624 cohorts through the system in sequence, as described directly above (see run in Table 1 for explanation). Our primary interest was to see how the IDBM behaved as a system over time. Thus, we operationalized system behavior as the prevalence of sexual initiation for each cohort when the cohort reaches its 8th tick (i.e., the final tick before exiting the system). Within a run, the concatenation of this measure across the 624 cohorts is a time series for which time is really cohort time (equally spaced measurements of the system, every 4 ticks, in Bin B). For all analyses below we transformed the raw data from the 624 cohorts into a moving average of prevalence. Each data point is therefore an average of the prevalence at tick 8 of 11 adjacent cohorts—including five before and after. By necessity, therefore, the first and last five cohort data points in a run are dropped, resulting in a total of 614 data points. We refer to these as cohorts, but keep in mind that the transformed cohorts each represent the average of 11 cohorts.

Bifurcation was defined as system behavior within a run such that the prevalence surpassed 0.32. Bifurcation was defined relative to the performance of the model within the parameters selected, and this value (of 0.32) was equivalent to +6 standard deviations from the moving average under conditions of no intercohort diffusion (results not presented). This was selected because of our interest in identifying extreme changes in system behavior. Using this definition, we computed two measurements for each run. First, mean proportion cohorts bifurcated represented the proportion of the 614 cohorts that reached bifurcation. This gives a sense of the overall amount of bifurcation in the system for any given run. Second, survival time represented the time (in cohort time) to the first bifurcation from the start of a run. This was either the time to bifurcation or, in the event of no bifurcation, was equal to 614. In effect, when a run did not bifurcate, it was right censored. This allowed us to use the standard tools of survival analysis to understand how this system operates with respect to bifurcation.

Agent social threshold function. Each agent could be in one of two states: sexually initiated or not. The state of an agent can change in one direction only, from not sexually initiated to initiated.

The social threshold function was: Change state if I $\geq$ T, where I is the amount of social influence in the agent's social context (defined below) and T is the agent's social threshold. Social threshold (T) was assigned to each agent in a cohort when the cohort entered the system, was drawn at random from a normal distribution $N(0.40, 0.12)$, bounded by (0,1), and never changed (see social threshold in Table 1).

If the agent was in Bin A, then I was dependent on whether the agent was linked to Bin B, as described in Equations 1 and 2,

$$I_{NotLinked} = \frac{N_{NeighborsInit}}{N_{NeighborsTotal}} \tag{1}$$

$$I_{Linked} = (1-\gamma)\frac{N_{NeighborsInit}}{N_{NeighborsTotal}} + \gamma P_{BinB} \tag{2}$$

where $N_{NeighborsInit}$ and $N_{NeighborsTotal}$ were the number of neighbors an agent had that were sexually initiated and the total number of neighbors, respectively. $P_{Bin\,B}$ represents the prevalence of sexual initiation in the cohort residing in Bin B and *gamma* (γ) signifies the strength of the influence from the Bin B cohort on the Bin A linked agents (γ ranged from 0.0 to 1.0) (see gamma in Table 1). As is evident from Equation 2, as γ increases the influence of the Bin B cohort increases relative to the influence of the agent's neighborhood. If the agent is in Bin B, then I was defined as in Equation 1.

Typical Behavior of the IDBM

Figure 2 presents the typical behavior of the IDBM. Each panel represents the system behavior of a single run. The parameters were fixed at identical values for

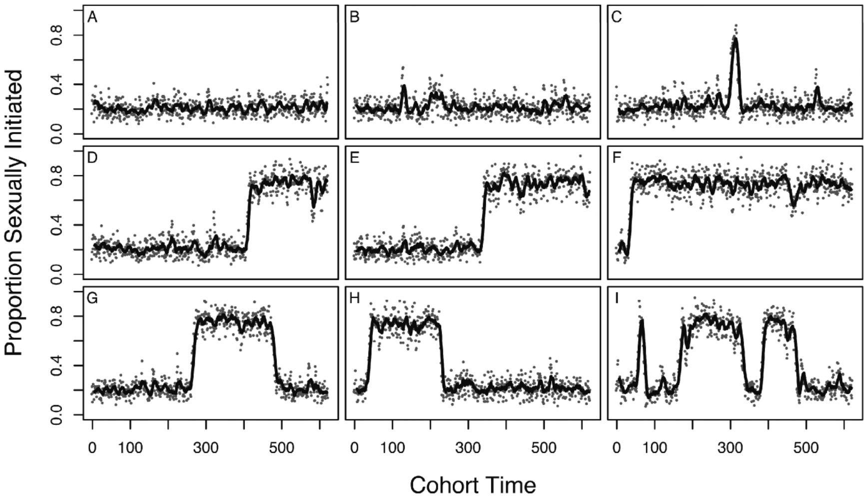

FIGURE 2 General behavior of intercohort diffusion of behavior model (IDBM). Presented here is a selection of nine IDBM simulations demonstrating the variety in possible system behaviors. The x-axes measure the simulation duration in cohort time. The y-axes depict the prevalence of each cohort that was sexually initiated in tick-age 8. The parameter values utilized to generate these figures were linked = 14%, $\gamma = .90$. These values were selected due to the particularly turbulent nature of simulations in this region in order to demonstrate several key behaviors of the IDBM. In particular the phenomena of bifurcation, or the sudden and sustained elevation in the prevalence of sexual initiation (Fig. 2D, 2E, 2F) and the sudden collapse of such bifurcated systems back to lower levels of prevalence (Fig. 2G, 2H, 2I).

each panel (parameters are described below). Figure 2 thus illustrates the variety of behaviors observed in the IDBM absent any parameter variation.

Individual data points in Figure 2 represent raw data (from 624 cohorts); the solid line is the moving average (614 cohorts). Figures 2A and 2B are simulations with relatively stable, lower levels of prevalence. The minor fluctuations in prevalence are due to variations in the social connectivity in each cohort. Figure 2C demonstrates that occasionally, in relatively stable systems, there are sudden, major spikes in prevalence. Figures 2D and 2E depict bifurcation. Simulations without intercohort diffusion never exhibit this feature (data not presented). Bifurcations are not predictable a priori and, as demonstrated in Figure 2F, they can occur at any point in the simulation, even immediately after the simulation begins. Once initiated, the diffusion from Bin B to Bin A cohorts bolsters the prevalence of sexual initiation in younger cohorts, thus sustaining the elevated levels. As demonstrated in Figures 2G, 2H, and 2I the elevated period is capable of spontaneous collapse back to lower levels.

Parameter Selection & Sensitivity Analysis

The IDBM is a theoretical, proof-of-concept model. Its sole purpose is to explore whether intercohort diffusion would give rise to systematic changes in prevalence. For this type of modeling effort, it is not essential to understand the system behavior across the full parameter space (there are seven parameters, see Table 1). Instead, the primary goal is to find a region of the parameter space that exhibits the system behavior in question, if possible, and then to explore key parameters in that region of the parameter space. This was also essential for developing and interpreting the simulated artificial interventions (see next section) because it provided an understanding of the baseline characteristics of the system.

We focused our efforts on the two parameters that control the strength of intercohort diffusion, the proportion of the younger cohort (Bin A) linked to the older (Bin B) and gamma (γ), the strength of the influence of the older cohort on the younger linked agents (see Equation 2). Thus, we ran a series of experiments crossing three levels of linked (12%, 14%, and 16%) with six levels of γ (0.0, 0.80, 0.85, 0.90, 0.95, and 1.0). For these experiments, we set the distribution of the social thresholds at mean 0.40 ($SD = .12$). Each cell of this 3×6 (linked by γ) design contained 50 experimental runs of 624 complete cohorts. Figure 3, rows A–C, shows the results for each level of linked. Within each row, the six values of γ are presented. The left panels show the mean proportion cohorts bifurcated where each bar represents the means of the 50 runs for proportion of cohorts bifurcated. The right panels supply Kaplan-Meier survival curves as an estimate of how the parameters affect time to first bifurcation. Each line represents the curve for the respective 50 runs at each level of γ.

It is clear that linked and γ affected the behavior of the system such that increasing either one increases the amount of bifurcation and decreases the survival probabilities of the system over time. However, the point of doing the parameter sweeps was to fix the values of linked and γ to conduct two demonstrations of the impact on the model of simple artificial interventions. Given the results above, we fixed linked = 14% and $\gamma = .90$. These values seem to provide a moderate amount of bifurcation and a survival curve that is midway between complete failure and complete survival.

After fixing the linked and γ parameters we also wanted to explore, in this parameter region, the effects of changing the mean of the social threshold distribution. This parameter will play a role in the artificial interventions so it was essential to test its sensitivity to changes at our given settings for linked and γ. Figure 3, row D, shows the results for each level of the mean social threshold from 0.38 to 0.42, using the same measures as were used for rows A–C. It was clear that this system, when fixed at linked = 14% and $\gamma = .90$, was highly sensitive to changes in the mean social threshold. 0.40 was the only value that provided much

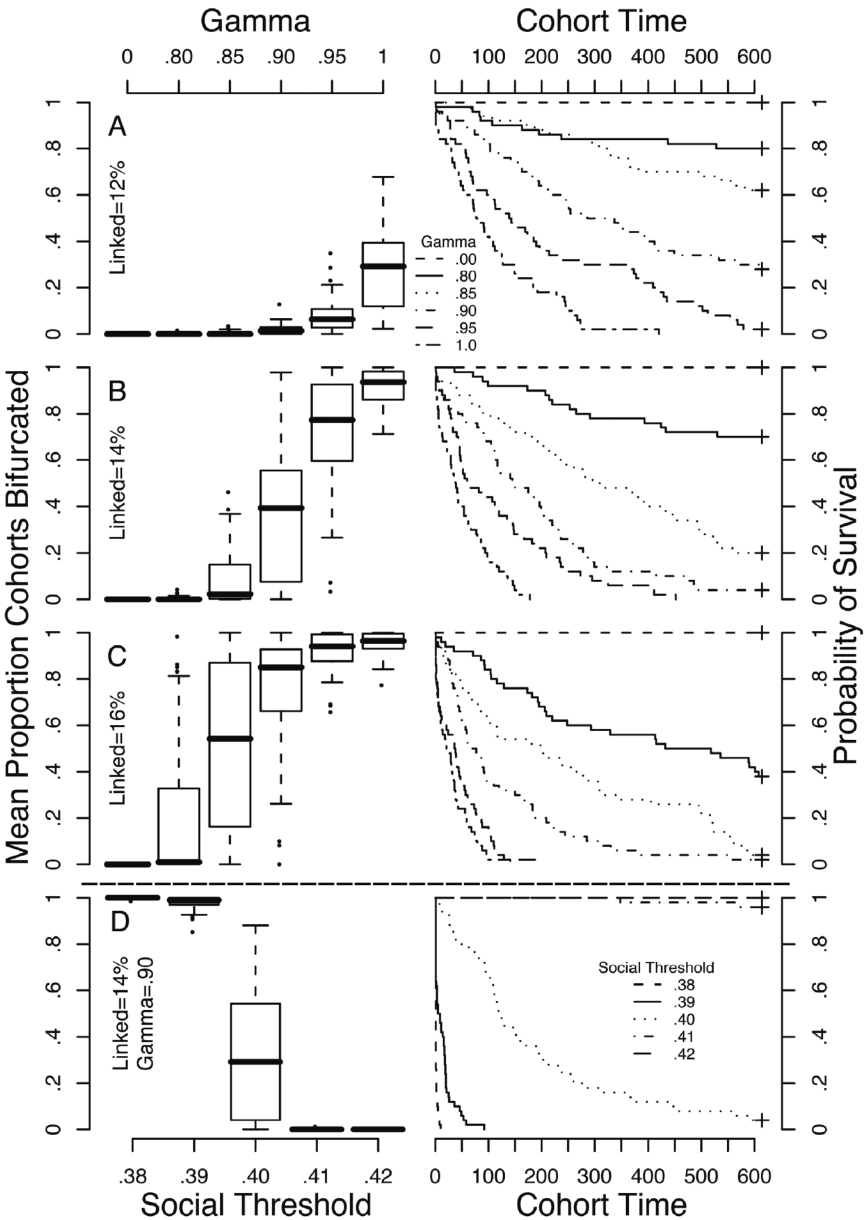

FIGURE 3 Parameter sweep results. Here we depict results from the parameter selection and calibration procedure. Rows A–C provide a comparison of the mean proportion cohorts bifurcated and probability of survival across each of the combinations of gamma (γ) and linked parameter values. Row D presents the results of a parameter sweep across mean social threshold values at fixed levels of linked (14%) and γ (0.90).

variability in the mean proportion of cohorts bifurcated or a survival curve that was not extreme (either very rapid decline in survival or all surviving). Therefore, we decided to use 0.40 as our social threshold for the remainder of our work.

Parameters that we did not explore were the number of agents (always fixed at 650), the proportion sexually initiated at model initialization (8%), the area of the physical environment, and the standard deviation of the social threshold distribution (0.12) (see Table 1 for discussion of each parameter).

Demonstrations of Artificial Interventions

The results presented below address two demonstrations where simple, artificial interventions targeted the social thresholds of the agents. The basic idea was to test whether system behavior (bifurcation) could be modified by manipulating the social thresholds. More specifically, we tested whether increasing the social thresholds of select agents (i.e., making agents less likely to initiate sexual behavior) would decrease the amount of bifurcation and the time to first bifurcation. It is critical to note that these interventions are a tool for testing system responses and are not intended to bear any great fidelity to real-world interventions.

Artificial Intervention 1. The first artificial intervention addressed the question of whether the age at intervention changes the system behavior. In other words, how early do you have to intervene to effect change in the system? Age in this case refers not to real-world age but to the stage of maturation (ticks 1–8) of agents in the system. We employed an intervention capable of shifting the entire distribution of the thresholds of agents in a cohort at a selected age by +0.02 (in effect this changed the average social threshold from 0.40 to 0.42). The increase of 0.02 was selected because, in the parameter selection experiments, runs with a mean social threshold of 0.42 did not exhibit any bifurcations (see Figure 3, Row D, mean social threshold = .42). So in effect this artificial intervention tests whether targeting agents at specific ages will have the same impact on system bifurcation behaviors. Three levels of intervention age were compared: intervention at tick 2, 4, and 6. Fifty runs of 624 cohorts were conducted for each intervention age for a total of 150 runs.

Artificial Intervention 2. Our second artificial intervention attempted to mimic, in principle, public health programs that are designed to target specific subpopulations. It builds on artificial intervention 1 by adding the constraint that either high- or low-risk individuals in the specified age groups are targeted. *High risk* was defined as agents with a social threshold of more than one standard deviation below the mean social threshold, *low risk* as more than one standard deviation above the mean social threshold. Otherwise, this intervention was identical to artificial intervention 1 with respect to the parameter values and the Intervention Age factor. This allows us to assess the joint and independent effects of Intervention Age and Risk Group on system behavior.

RESULTS

The results for the artificial interventions are presented in the same format as the parameter calibrations in Figure 3. Each intervention was analyzed separately.

Artificial Intervention 1

The results for artificial intervention 1 are presented in Figure 4. Panel A shows the mean proportion cohorts bifurcated for each intervention age. There is a clear increase in bifurcations as age increases. Panel B shows the survival curves by intervention age. Again, there is a clear effect. When intervening at age 2, the model does not exhibit any bifurcations. Increasing the intervention age, however, does begin to affect the survival curves. To test this formally, we used a Cox proportional hazards model incorporating intervention age as the predictor. The equation for this model asserts that the hazard rate for the ith run in the data is

$$h(t|x_i) = h_0(t)e^{(x_i\beta_x)} \tag{3}$$

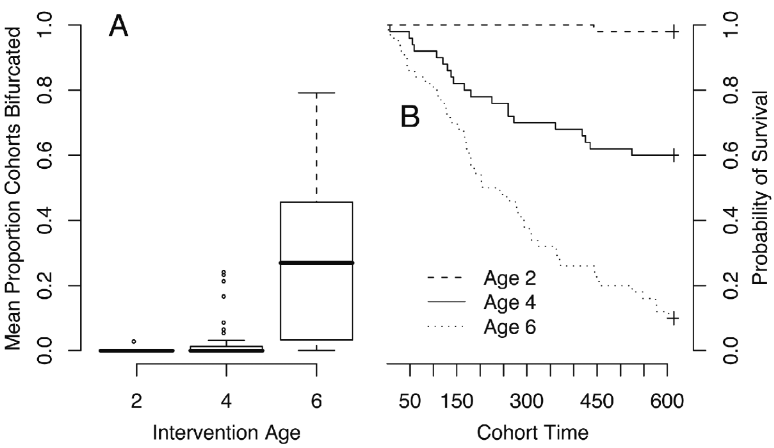

FIGURE 4 This figure presents results from the Artificial Intervention 1. Fig. 4A compares the mean proportion cohorts bifurcated for each intervention age targeted (2, 4, and 6). For each Intervention Age, 50 simulations were run. Fig. 4B compares the survival time to first bifurcation for each intervention age.

where x_i is the categorical levels of intervention age and the β_x is the coefficient estimated from the data. In these data, each data point is one run of 614 cohorts (the moving average was used, not raw prevalence); 50 runs per intervention age (3 levels Intervention Age $\times$ 50 runs $=$ 150 runs). Furthermore, runs that never bifurcated were right censored and given a survival time of 614 (in cohort time). The results are given in exponentiated coefficients to represent the difference in the proportional hazard between levels of intervention age (age 2 was the reference category). Runs of Intervention Age 4 were 26.9 times more likely to bifurcate than Intervention Age 2 (95% CI [3.49, 206.44], $p < 0.01$); Age 6 was 97.7 times more likely (95% CI [12.94, 735.10, $p < 0.0001$).

Artificial Intervention 2

The results for artificial intervention 2 are presented in Figure 5. Panels A and B show the results for the low-risk intervention. It is clear that, when targeting this subpopulation, there was no effect at any intervention age, either in increases in the mean proportion cohorts bifurcated or in the survival curves. Panels C and D show the results for the high-risk intervention. Here, there was a clear effect for Intervention Age. When intervening at an earlier age, the mean proportion cohorts bifurcated decreased and the survival curves showed longer times to bifurcation. Taking Figure 5 as a whole suggests an interaction between Intervention Age and Risk Group. To test this formally, we compared two separate Cox proportional hazards models. For Model 1, x_i and the associated β_x represent the categorical levels of Intervention Age and Risk Group (age 2 and high risk were the reference categories). For Model 2, x_i and the associated β_x represent the categorical levels of intervention age and risk group (age 2 and high risk were the reference categories) and the two interaction terms (Intervention Age 4 $\times$ Low Risk and Intervention Age 6 $\times$ Low Risk). The general form of the equation is identical to Equation 3 as were the presentation of the results. The structure of these data was identical to that for artificial intervention 1. Table 2 presents the results. Model 1 showed a strong effect for both Intervention Age and Risk Group. As intervention age increased, so did the likelihood of bifurcation. Furthermore, the low-risk runs were less likely to bifurcate than the high-risk runs. Model 1 is nested in Model 2. Therefore, we computed an analysis of deviance test to determine whether adding the Intervention Age $\times$ Risk Group interaction added significantly to the model. The residual deviances of Model 1 and Model 2 were 2475.27 ($df = 297$) and 2454.32 ($df = 295$), respectively. The analysis of deviance test was significant ($p < 0.0001$, $df = 2, 295$) indicating the superiority of Model 2. By adding the interaction term, Intervention Age did not remain predictive of likelihood of bifurcation whereas Risk Group did and kept the same relative direction (reducing the likelihood of bifurcation). However, this result was qualified by a significant Intervention Age $\times$ Risk Group interaction. Figure 5 illustrates the pattern

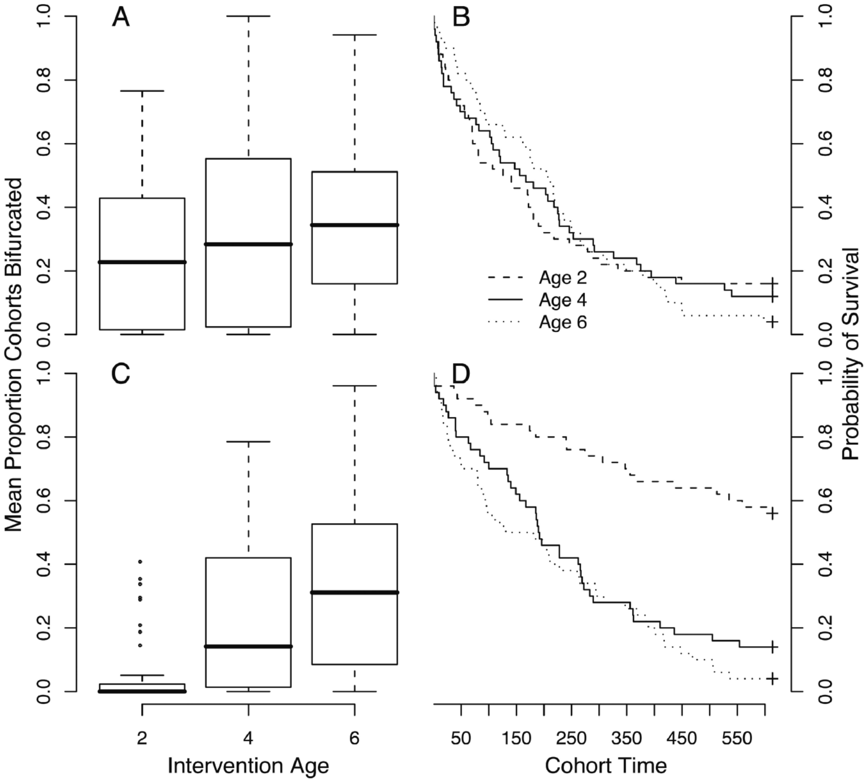

FIGURE 5 This figure presents results from Artificial Intervention 2. Fig. 5A and 5B depict the mean proportion cohorts bifurcated and the survival time to first bifurcation, respectively, for interventions targeting low-risk agents. Fig. 5C and 5D depict the same for interventions targeting high-risk agents. *High risk* was defined as agents with a social threshold of more than one standard deviation below the mean social threshold and *low risk* as agents with social thresholds more than one standard deviation above the mean social threshold. Within each level of *risk group* is a comparison of the impact on system behavior across the three intervention ages (2, 4, and 6). For each Intervention Age, 50 simulations were run.

found in the coefficients. In conclusion, Intervention Age only had an effect for the high-risk group.

DISCUSSION

In qualitative terms, the IDBM matched the empirical regularities found in long-term trends in sexual initiation prevalence in the United States—for example,

TABLE 2
Cox Proportional Hazard Results for Artifical Intervention 2

	exp(coef)	*95% C.I.*	$p < x$
Model 1			
Age 4	1.738	1.254 – 2.411	0.001
Age 6	1.989	1.438 – 2.750	0.0001
Low Risk	0.672	0.521 – 0.867	0.01
Model 2			
Age 4	0.978	0.641 – 1.493	NS
Age 6	1.044	0.689 – 1.582	NS
Low Risk	0.246	0.147 – 0.414	0.0001
Age 4 × Low Risk	3.506	1.797 – 6.838	0.001
Age 6 × Low Risk	4.161	2.162 – 8.010	0.0001

systematic increases, decreases, and equilibria (see Figure 2). This is somewhat surprising given that the IDBM is restricted to social diffusion as its sole mechanism for change. Furthermore, the artificial interventions had a measureable effect on the behavior of the system in an interpretable way. Intervening at a younger age clearly reduced bifurcations. However, Artificial Intervention 2 showed that the efficacy of intervening at a younger age worked only if the high-risk group was targeted.

At face value, the behavior of the IDBM was severe—bifurcations were sharp, and the sensitivity to parameter values was high. This is to be expected in a complex system. However, the most important question is, can we infer that these characteristic behaviors say anything about long-term changes in human sexual initiation prevalence in the real world? In other words, is human sexual initiation governed by self-organization and characterized by bifurcations? At this point, provisionally, we can only say that the qualitative behavior of the IDBM matches, in an abstract way, what we see in the empirical trends in sexual initiation. Next, we recommend modifications to the IDBM that would afford more justified inferences to the real world.

There are three key omissions in the structure of the IDBM that should be recognized. First, though a threshold will theoretically incorporate individual-level characteristics implicitly (Granovetter, 1978), we did not imbue agents with individual-level characteristics that are known risk factors for sexual initiation in the IDBM (e.g., personality, age, gender, race/ethnicity, attitudes toward sex, intentions, and parental monitoring). Second, we did not incorporate exogenous influences, such as messages in mixed media, or changes in policy and birth control technologies, into the IDBM. Third, the social network structure in the IDBM uses random spatial distributions on a fixed Cartesian grid to approximate true network structures. A related issue is that the network density of the

IDBM is greater than empirical data from school friendship networks would indicate (Moody, 1999). Future work should incorporate realistic network structure because it could affect the diffusion dynamics.

Two crucial next steps are (1) to increase the fidelity of the model to the real world, particularly in the three areas indicated, so as to boost the validity of inferences from the model to the real world, and (2) to collect data that is isomorphic to the structure of the IDBM (e.g., a single grade in one school). This will bring the IDBM, and associated inferences to human sexual initiation, more validity.

Figure 2 demonstrates how much change in the system is possible and how it is very difficult to predict, even when parameter values are held constant. Every slice of cohort time in Figure 2 is identical with respect to the allowable parameters, but system-level prevalence changes dramatically. The only difference in time slices is that the developmental context has changed. This concept, we argue, is very important for developmental science. Complex systems approaches, such as the ABM technique illustrated here, are excellent tools for understanding developmental context. The IDBM reveals clearly how outcomes for agents—individually and in aggregate—are largely dependent on the structure of the immediate social context and the larger historical context they exist within. The study of individual-level development can be strengthened through the addition of complex systems approaches and perspectives.

ACKNOWLEDGMENTS

This work was funded, in part, by a seed grant from The Institute for Social and Economic Research in Policy at Columbia University.

REFERENCES

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: W. H. Freeman.

Bearman, P. S., & Moody, J. (2004). Suicide and friendships among American adolescents. *American Journal of Public Health, 94*(1), 89–95.

Brown, J. D. (2002). Mass media influences on sexuality. *Journal of Sex Research, 39*(1), 42–45.

Centers for Disease Control and Prevention. (2010). *Trends in current cigarette smoking among high school students and adults, United States, 1965–2009*. Retrieved from www.cdc.gov/tobacco/data_statistics/tables/trends/cig_smoking/index.htm

Christakis, N. A., & Fowler, J. H. (2008). The collective dynamics of smoking in a large social network. *New England Journal of Medicine, 358*(21), 2249–2258.

Coontz, S. (1992). *The way we never were: American families and the nostalgia trap*. New York, NY: Basic Books.

Finer, L. (2007, Jan.-Feb.). Trends in premarital sex in the United States, 1954–2003. *Public Health Reports, 122*, 73–78.

Fowler, J. H., & Christakis, N. A. (2010). Cooperative behavior cascades in human social networks. *Proceedings of the National Academy of Sciences USA, 107*(12), 5334–5338.

Galea, S., Hall, C., & Kaplan, G. A. (2009). Social epidemiology and complex system dynamic modelling as applied to health behavior and drug use research. *International Journal of Drug Policy, 20*, 209–216.

Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology, 83*(6), 1420–1443.

Hill, A. L., Rand, D. G., Nowak, M. A., & Christakis, N. A. (2010a). Emotions as infectious disease in a large social network: the SISa model. *Proceedings of the Royal Society: Biology*, 1–19.

Hill, A. L., Rand, D. G., Nowak, M. A., & Christakis, N. A. (2010b). Infectious disease modeling of social contagion in networks. *PLoS Computational Biology, 6*(11), 1–15.

Link, B. G. (2008, December). Epidemiological sociology and the social shaping of population health. *Journal of Health and Social Behavior, 29*, 367–384.

Miller, J. H., & Page, S. E. (2007). *Complex adaptive systems: An introduction to computational models of social life*. Princeton, NJ: Princeton University Press.

Moody, J. W. (1999). *The structure of adolescent social relations: Modeling friendship in dynamic social settings* (Unpublished doctoral DISSERTATION). Retrieved from ProQuest Dissertations and Theses database. (UMI No. AAT 9954682).

Newman, M. (2010). *Social networks: An introduction*. Oxford, UK: Oxford University Press.

Rahmandad, H., & Sterman, J. (2008). Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. *Management Science, 54*(5), 998–1014.

Read, S. J., Monroe, B. M., Brownstein, A. L., Yang, Y., Chopra, G., & Miller, L. C. (2010). A neural network model of the structure and dynamics of human personality. *Psychological Review, 117*(1), 61–92.

Rodgers, J. L., Rowe, D. C., & Buster, M. (1998). Social contagion, adolescent sexual behavior, and pregnancy: A nonlinear dynamic EMOSA model. *Developmental Psychology, 34*(5), 1096–1113.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York, NY: Free Press.

Romer, D., & Stanton, B. F. (2003). Feelings about risk and the epidemic diffusion of adolescent sexual behavior. *Prevention Science, 4*(1), 39–53.

Santelli, J. A., Orr, M. G., Lindberg, L., & Diaz, D. C. (2009). Changing behavioral risk for pregnancy among high school students in the United States, 1991–2007. *Journal of Adolescent Health, 45*(1), 25–32.

Shoda, Y., LeeTiernan, S., & Mischel, W. (2002). Personality as a dynamical system: Emergence of stability and distinctiveness from intra- and interpersonal interactions. *Personality and Social Psychology Review, 6*(4), 316–325.

Singh, S., & Darroch, J. E. (1999). Trends in sexual activity among adolescent American women: 1982–1995. *Family Planning Perspectives, 31*(5), 212–219.

Smith, K. P., & Christakis, N. A. (2008). Social networks and health. *Annual Review of Sociology, 34*, 405–429.

Ueno, K. (2005). The effects of friendship networks on adolescent depressive symptoms. *Social Science Research, 34*(3), 484–510.

Valente, T. W. (1995). *Network models of the diffusion of innovations*. Cresskill, NJ: Hampton Press.

Valente, T. W. (2010). *Social networks and health: Models, methods, and applications*. Oxford, UK: Oxford University Press.

Watts, D. J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences of the United States of America, 99*(9), 5766–5771.