Heterogeneity and Network Structure in the Dynamics of Diffusion: Comparing Agent-Based and Differential Equation Models

Hazhir Rahmandad

Department of Industrial and Systems Engineering, Virginia Tech, Falls Church, Virginia 22043, hazhir@vt.edu

John Sterman

Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142,
jsterman@mit.edu

When is it better to use agent-based (AB) models, and when should differential equation (DE) models be used? Whereas DE models assume homogeneity and perfect mixing within compartments, AB models can capture heterogeneity across individuals and in the network of interactions among them. AB models relax aggregation assumptions, but entail computational and cognitive costs that may limit sensitivity analysis and model scope. Because resources are limited, the costs and benefits of such disaggregation should guide the choice of models for policy analysis. Using contagious disease as an example, we contrast the dynamics of a stochastic AB model with those of the analogous deterministic compartment DE model. We examine the impact of individual heterogeneity and different network topologies, including fully connected, random, Watts-Strogatz small world, scale-free, and lattice networks. Obviously, deterministic models yield a single trajectory for each parameter set, while stochastic models yield a distribution of outcomes. More interestingly, the DE and mean AB dynamics differ for several metrics relevant to public health, including diffusion speed, peak load on health services infrastructure, and total disease burden. The response of the models to policies can also differ even when their base case behavior is similar. In some conditions, however, these differences in means are small compared to variability caused by stochastic events, parameter uncertainty, and model boundary. We discuss implications for the choice among model types, focusing on policy design. The results apply beyond epidemiology: from innovation adoption to financial panics, many important social phenomena involve analogous processes of diffusion and social contagion.

Key words: agent-based models; networks; scale free; small world; heterogeneity; epidemiology; simulation; system dynamics; complex adaptive systems; SEIR model
History: Accepted by Wallace J. Hopp, stochastic models and simulation; received November 2, 2004. This paper was with the authors 1 year and $11\frac{1}{2}$ months for 4 revisions. Published online in *Articles in Advance* March 18, 2008.

Introduction

Spurred by growing computational power, agent-based (AB) modeling is increasingly applied to physical, biological, social, and economic problems previously modeled with nonlinear differential equations (DEs). Both approaches have yielded important insights. In the social sciences, agent models explore phenomena from the emergence of segregation to organizational evolution to market dynamics (Schelling 1978, Levinthal and March 1981, Carley 1992, Axelrod 1997, Lomi and Larsen 2001, Axtell et al. 2002, Epstein 2006, Tesfatsion 2002). Differential and difference equation models, also known as compartmental models, have an even longer history in social science, including innovation diffusion (Mahajan et al. 2000) and epidemiology (Anderson and May 1991).

When should AB models be used, and when are DE models appropriate? Each method has strengths and weaknesses. The importance of each depends on the model purpose. Nonlinear DE models can easily encompass a wide range of feedback effects, but typically aggregate agents into a relatively small number of states (compartments). For example, innovation diffusion models may aggregate the population into categories including unaware, aware, in the market, adopters, and so on (Urban et al. 1990, Mahajan et al. 2000). However, within each compartment, people are assumed to be homogeneous and well mixed; the transitions among states are modeled as their expected value, possibly perturbed by random events. In contrast, AB models can readily include heterogeneity in individual attributes and in the network structure of their interactions; like DE models,

AB models can be deterministic or stochastic and can capture feedback effects.

The granularity of AB models comes at some cost. First, the extra complexity significantly increases computational requirements, constraining the ability to conduct sensitivity analysis. A second cost of agent-level detail is the cognitive burden of understanding model behavior. Linking the behavior of a model to its structure becomes more difficult as model complexity grows. Finally, limited time and resources force modelers to trade off disaggregate detail and the breadth of the model boundary. Here model boundary stands for the richness of the feedback structure captured endogenously in the model (Meadows and Robinson 1985, Sterman 2000). For example, an agent-based demographic model may portray each individual separately, but assume exogenous fertility and mortality; such a model has a narrow boundary. In contrast, an aggregate model may lump the entire population into a single compartment, but model fertility and mortality as functions of food per capita, health care, pollution, norms for family size, etc., each of which, in turn, are modeled endogenously; such a model has a broad boundary. DE and AB models may in principle fall anywhere on these dimensions of disaggregation and scope. In particular, there is no intrinsic limitation that prevents AB models from incorporating behavioral feedback effects or encompassing a broad model boundary. In practice, however, where time, budget, and computational resources are limited, modelers must trade off disaggregate detail and breadth of boundary. Choosing wisely is central in selecting appropriate methods for any problem.

The stakes are large. Consider potential bioterror attacks. Kaplan et al. (2002) used a deterministic nonlinear DE model to examine smallpox attack in a large city, comparing mass vaccination (MV), in which essentially all people are vaccinated after an attack, to targeted vaccination (TV), in which health officials trace and immunize those contacted by potentially infectious individuals. Capturing vaccination capacity and logistics explicitly, they conclude MV significantly reduces casualties relative to TV. In contrast, using different AB models, Eubank et al. (2004) and Halloran et al. (2002) conclude TV is superior, while Epstein et al. (2004) favor a hybrid strategy. The many differences among these models make it difficult to determine whether the conflicting conclusions arise from relaxing the perfect mixing and homogeneity assumptions of the DE (as argued by Halloran et al. 2002) or from other assumptions such as the size of the population (ranging from 10 million for the DE model to 2,000 in Halloran et al. 2002 to 800 in Epstein et al. 2004), other parameters, or boundary differences such as whether capacity constraints on immunization are included (Koopman 2002, Ferguson et al.

2003, Kaplan and Wein 2003). Kaplan and Wein (2003) and Kaplan et al. (2003) show that their DE model closely replicates the Halloran et al. (2002) AB results when simulated with the Halloran et al. parameters, including vaccination rates, population, and initial attack size, concluding that parameterization accounts for the different conclusions, not differences in mixing and homogeneity.

Here we carry out controlled experiments to compare AB and DE models in the context of contagious disease. We choose disease diffusion for four reasons. First, the dynamics of contagion involve important characteristics of complex systems, including positive and negative feedbacks, time delays, nonlinearities, stochastic events, and individual heterogeneity. Second, network topologies linking individuals are important in the diffusion process (Davis 1991, Watts and Strogatz 1998, Barabasi 2002, Rogers 2003), providing a strong test for differences between the two approaches. Third, the DE paradigm is well developed in epidemiology (for reviews see Anderson and May 1991 and Hethcote 2000); AB models also have a long history (e.g., Abbey 1952) and have recently gained momentum (for reviews see Newman 2002, 2003 and Watts 2004).

Finally, diffusion is a fundamental process in diverse physical, biological, social, and economic settings. Many diffusion phenomena in human systems involve processes of social contagion analogous to infectious disease, including word of mouth, imitation, and network externalities. From the diffusion of innovations to rumors, financial panics, and riots, contagion-like dynamics, and formal models of them, have a rich history in the social sciences (Bass 1969, Watts and Strogatz 1998, Mahajan et al. 2000, Barabasi 2002, Rogers 2003). Insights into the advantages and disadvantages of AB and DE models in epidemiology can inform understanding of diffusion in many domains of concern to social scientists and managers.

Our procedure is as follows. We develop a stochastic AB version of the classic SEIR model, a widely used nonlinear deterministic DE model. The DE divides the population into four compartments: susceptible (S), exposed (E), infected (I), and removed (R). In the AB model, each individual is separately represented and must be in one of these four states. Both the AB and DE models use the same parameters. Therefore, any differences in outcomes arise only from relaxing the restrictive assumptions of the DE model. In practice, DE modelers add compartments to capture heterogeneity in individuals and their contact networks, for example, disaggregating by biological or behavioral attributes (e.g., differences in age or contact frequencies), or by location (as in patch models; see, e.g., Riley 2007). Here we use the classic SEIR model to maximize potential differences between the

two approaches. We run the AB model under five widely used network topologies (fully connected, random, small world, scale-free, and lattice) and test each with homogeneous and heterogeneous individuals. We compare the resulting diffusion dynamics on a variety of metrics relevant to public health, including cumulative cases, peak prevalence, and the speed the disease spreads (the time available for health officials to respond).

The most obvious difference between the models we compare is that, for given parameters, the stochastic AB model generates a distribution of outcomes, while the deterministic DE generates a single path representing the expected trajectory under the mean-field approximation for contacts between infectious and susceptible individuals. Further, due to chance events, the epidemic never takes off in some realizations of the stochastic model. Deterministic models, whether DE or AB, cannot generate this mode of behavior. Capturing outcome variability in the DE paradigm requires moving to a stochastic compartment model, an intermediate method between deterministic models and the full stochastic AB representation. More interesting are differences due to network topology and individual heterogeneity. On average, diffusion slows as contact networks become more tightly clustered compared to the DE. On average, heterogeneity accelerates the initial take-off, as highly connected individuals quickly spread the disease, but reduces overall diffusion as these same individuals quickly exit the infectious pool.

In a second set of tests, we also examine the ability of the DE model to capture the dynamics of each network structure in the realistic situation where parameters are poorly constrained by biological and clinical data. Epidemiologists often estimate potential diffusion, for both novel and established pathogens, by fitting models to the aggregate data as an outbreak unfolds (Dye and Gay 2003, Lipsitch et al. 2003, Riley et al. 2003). Calibration of innovation diffusion and new product marketing models is similar (Mahajan et al. 2000). We mimic this practice by treating the AB simulations as the "real world" and fitting the DE model to them. On average, the fitted models closely match the individual AB realizations under all network topologies and heterogeneity conditions tested. However, the estimated parameters are biased in the highly clustered and heterogeneous cases. Further, the ability to fit such data does not imply that the AB and calibrated DE models will respond to policy interventions in the same way, demanding caution in their use. When different models yield different inferences about policies, it is important to identify the assumptions responsible to guide data collection, to improve the models, and to select the most appropriate model for the purpose at hand.

The implications of the differences across models depend on the purpose of the analysis. Here we focus on the policy context. Policymakers face a world of time pressure, inadequate data, and limited knowledge of parameters such as pathogen virulence, transmissibility, incubation latency, treatment efficacy, etc. Further, the appropriate boundary for analysis is often unclear: resources for vaccination and treatment may be limited; an outbreak, whether natural or triggered by bioterror, may alter the behavior of the public and first-responders, endogenously disrupting the contact networks that feed back to condition disease spread through processes of risk amplification and attenuation (Kasperson et al. 1988, Glass and Schoch-Spana 2002). Hence, we consider whether the differences in mean behavior between DE and AB models are large relative to the uncertainties policymakers face. We also consider how these differences in mean behavior might affect the assessment of the costs and benefits, and hence the optimality, of policies.

The mean behavior of different models may be significantly different in the statistical sense, yet be small relative to the variation in output caused by uncertainty about parameters, model boundary, and stochastic events (McCloskey and Ziliak 1996). For example, consider the variability in outcomes generated by a stochastic AB model. Each realization of the model will differ: some exhibiting fast diffusion, some slow; some with many individuals afflicted, some with fewer, depending on the chance nature of contacts between infectious and susceptible individuals. An ensemble of many simulations generates the distribution of possible epidemics, but only one will be observed in a particular outbreak. Several questions may now be asked.

One important question is whether the expected values of key metrics differ in different models. For example, does the mean value of peak prevalence under a scale-free network differ from the value generated by the corresponding deterministic compartment model? By running a sufficiently large number of simulations, sampling error can be made arbitrarily small, and any differences in the mean behavior of the models will be highly statistically significant.

Another question is whether the differences among means are significant from the point of view of policymakers seeking appropriate responses to a potential outbreak. Models with similar "base-case" behavior can have similar or different responses to policies, and, conversely, models with different base behaviors may nevertheless yield the same inferences about policy impacts. Differences in policy response across models can be statistically significant, yet small relative to uncertainty in parameters, network structure, individual attributes, and model boundary. Policymakers must assess the practical significance of

each model assumption given the likely range of outcomes generated by all sources of uncertainty, not only uncertainty caused by random events.

The results document a number of differences between the DE and mean AB dynamics. The results, for both the base-case and calibrated DE models, also show that the differences between the deterministic compartment model, with its assumptions of homogeneous individuals and perfect mixing, and the mean behavior of the stochastic AB models are often small compared to the variability in AB outcomes caused by chance encounters among individuals, at least for the public health metrics examined here. However, cost/benefit assessments of policy interventions, and hence the optimal policy, can depend on network structure and model boundary, underscoring the importance of sensitivity analysis across these dimensions.

The next section reviews the literature comparing AB and DE approaches. We then describe the models, the design of the simulation experiments, and results, closing with implications and directions for future work.

A Spectrum of Aggregation Assumptions

AB and DE models should be viewed as regions in a space of modeling assumptions, not as incompatible modeling paradigms. Aggregation is one dimension of that space. Models can range from lumped deterministic differential equations (also called deterministic compartmental models), to stochastic compartmental models, in which the continuous variables of the DE are replaced by counts of discrete individuals, to event history models, where the states of individuals are tracked but their network of relationships is ignored, to models with explicit contact networks linking individuals (e.g., Koopman et al. 2001, Riley 2007).

A few studies compare AB and DE models. Axtell et al. (1996) call for "model alignment" or "docking" and illustrate with the Sugarscape model. Edwards et al. (2003) contrast an AB model of innovation diffusion with an aggregate model, finding that the two can diverge when multiple attractors exist in the deterministic model. In epidemiology, Jacquez and O'Neill (1991) and Jacquez and Simon (1993) analyze the effects of stochasticity in individual-level susceptible-infectious-susceptible (SIS) and susceptible-infectious (SI) models, finding some differences in mean behavior for small populations. However, the differences practically vanish for homogeneous populations above 100. Similarly, Gani and Yakowitz (1995) examine deterministic approximations to stochastic disease diffusion processes, and

find a high correspondence between the two for larger populations. Greenhalgh and Lewis (2001) compare a stochastic model with the deterministic DE version in the case of AIDS spread through needle sharing, and find similar behavior for those cases in which the epidemic takes off.

Heterogeneity has also been explored in models with different mixing sites for population subgroups. Anderson and May (1991, Chap. 12) show that the immunization fraction required to quench an epidemic rises with heterogeneity if immunization is implemented uniformly, but falls if those most likely to transmit the pathogen are the focus of immunization. Ball et al. (1997) and Koopman et al. (2002) find expressions for cumulative cases and epidemic thresholds in stochastic susceptible-infectious-recovered (SIR) and SIS models with global and local contacts, finding that the behavior of deterministic and stochastic DE models can diverge for small populations, low basic reproduction rates (R_0), or highly clustered contact networks where transmission occurs in mixing sites such as schools and offices. Keeling (1999) formulates a DE model that approximates the effects of spatial structure when networks are highly clustered. Chen et al. (2004) develop AB models of smallpox, finding the dynamics generally consistent with DE models. In sum, AB and DE models of the same phenomenon sometimes agree and sometimes diverge, especially when compartments contain smaller populations. Multiple network topologies and heterogeneity conditions have not been compared, and the practical significance of differences in mean behavior relative to uncertainties in stochastic events, parameters, and model boundary has not been explored.

Model Structure

The SEIR model is a deterministic nonlinear differential equation model in which all members of a population are in one of four states—susceptible, exposed, infected, or removed. Contagious individuals can infect susceptibles before they are "removed" (i.e., recover or die). The exposed compartment captures latency between infection and the emergence of symptoms. Depending on the disease, exposed individuals may become infectious before symptoms emerge, and can be called early-stage infectious. Typically, such individuals have more contacts than those in later stages because they are asymptomatic.

SEIR models have been successfully applied to many diseases. Additional compartments are often introduced to capture more complex disease lifecycles, diagnostic categories, therapeutic protocols, population heterogeneity and mixing patterns, birth or recruitment of new susceptibles, loss of immunity,

etc. (see Anderson and May 1991, Murray 2002). In this study, we maintain the standard assumptions of the classic SEIR model (four stages, fixed population, permanent immunity). The DE implementation of the model imposes several additional assumptions, including perfect mixing and homogeneity of individuals within each compartment and mean-field aggregation (the flows between compartments equal the expected value of the sum of the underlying probabilistic rates for each individual). To derive the differential equations, consider the rate at which each infectious individual generates new cases:

$$c_{is} * \text{Prob}(\textit{contact with susceptible})$$
$$* \text{Prob}(\textit{transmission/contact with susceptible}), \quad (1)$$

where the contact frequency c_{is} is the expected number of contacts between infectious individual i and susceptible individual s; homogeneity implies c_{is} is a constant, denoted c_{IS}, for all individuals i, s. If the population is well mixed, the probability of contacting a susceptible individual is simply the proportion of susceptibles in the total population, S/N. Denoting the probability of transmission given contact between individuals i and s, or infectivity, as i_{is} (which, under homogeneity, equals i_{IS} for all i, s), and summing over the infectious population yields the total flow of new cases generated by contacts between the I and S populations, $c_{IS} * i_{IS} * I * (S/N)$. The number of new cases generated by contacts between the exposed and susceptibles is formulated analogously, yielding the total infection rate, f:

$$f = (c_{ES} i_{ES} E + c_{IS} i_{IS} I)(S/N). \quad (2)$$

To model emergence and recovery, consider these to be Markov processes with certain transition probabilities. In the classic SEIR model, each compartment is assumed to be well mixed so that the probability of emergence (or recovery) is independent of how long an individual has been in the E (or I) state. Denoting the individual hazard rates for emergence and recovery as ε and δ, the mean emergence time and disease duration are then $1/\varepsilon$ and $1/\delta$, respectively. Summing over the E and I populations and taking expected values yields the flows of emergence and recovery:

$$e = \varepsilon E \quad \text{and} \quad r = \delta I. \quad (3)$$

The full model is thus

$$\frac{dS}{dt} = -f, \quad \frac{dE}{dt} = f - e, \quad \frac{dI}{dt} = e - r, \quad \frac{dR}{dt} = r. \quad (4)$$

Equation (3) implies the probabilities of emergence and recovery are independent of how long an individual has been in the E or I states, respectively, and results in exponential distributions for the residence times in these states. Exponential residence times are not realistic for most diseases, where the probability of emergence and recovery is initially low, then rises, peaks, and falls. Note, however, that any lag distribution can be captured through the use of partial differential equations, approximated in the ordinary differential equation (ODE) paradigm by adding additional compartments within the exposed and infectious categories (Jacquez and Simon 2002). For simplicity, we maintain the assumption of a single compartment per disease stage of the classic SEIR model.

The AB model relaxes the perfect mixing and homogeneity assumptions of the DE. Each individual $j \in (1, \ldots, N)$ is in one of the four states S, E, I, or R. The individual state transitions $f[j], e[j]$, and $r[j]$ equal one at the moment of infection, emergence, and recovery, respectively, and zero otherwise, and depend on individual attributes such as contact frequencies and on the chances of interaction with others as specified by the contact network. The aggregate flows f, e, and r over any interval dt are the sum of the individual transitions during that interval. The online supplement (provided in the e-companion)[1] details the formulation of the AB model and shows how the DE model can be derived from it by assuming homogeneous agents and applying the mean-field approximation.

A central parameter in epidemic models is the basic reproduction number, R_0, the expected number of new cases each contagious individual generates before removal, assuming all others are susceptible. The base-case parameters yield $R_0 = 4.125$ (Table 1), similar to diseases like smallpox, $R_0 \approx 3$–6 (Gani and Leach 2001), and SARS, $R_0 \approx 2$–7 (Lipsitch et al. 2003, Riley et al. 2003). The base value provides a good opportunity to observe potential differences between DE and AB models: diseases with $R_0 < 1$ pose little risk to public health, while those with $R_0 \gg 1$, e.g., measles, cause a severe epidemic in (nearly) any network. The AB models use the same infectivities and expected residence times, and we choose individual contact frequencies so that mean total contact rates in each network and heterogeneity condition are the same as the DE model. We set the population $N = 200$, all susceptible except for two randomly chosen exposed individuals. Although small compared to settings of interest in policy design, e.g., cities, the effects of random events and network type are likely to be more pronounced in small populations (Gani and Yakowitz 1995), providing a stronger test for differences between the DE and AB models. A small population also reduces computation time, allowing more

Table 1   Base-Case Parameters

Parameter (dimensionless)			Parameter			Units
Infectivity, exposed	i_{ES}	0.05	Total population	N	200	Person
Infectivity, infectious	i_{IS}	0.06	Contact rate, exposed	c_{ES}	4	1/Day
Basic reproduction rate	R_0	4.125	Contact rate, infectious	c_{IS}	1.25	1/Day
Average links per node	k	10	Average incubation time	$1/\varepsilon$	15	Day
Prob. of long-range links (SW)	p_{sw}	0.05	Average duration of illness	$1/\delta$	15	Day
Scaling exponent (scale-free)	γ	2.60				

Note. The online supplement provides full details for the AB simulations.

extensive sensitivity analysis. The DE has four state variables; computation time is negligible for all N. The AB model has $4N$ state variables and must also track interactions among the N individuals, implying that computation time can grow at rates up to $O(N^2)$, depending on the contact network. We report sensitivity analysis of R_0 and N below. The online supplement includes the models and full documentation.

Experimental Design
We vary both the network structure of contacts among individuals and the degree of individual heterogeneity in the AB model, and compare the resulting dynamics to the DE. We implement a full factorial design with five network structures and two heterogeneity conditions. In each of the 10 conditions, we generate an ensemble of 1,000 simulations of the AB model, each with different realizations of the random variables determining contacts, emergence, and recovery. Because the parameters in each simulation are identical to the DE model, differences in outcomes can only be due to differences in network topology, heterogeneity among individuals, or the discrete, stochastic treatment of individuals in the AB model.

Network Topology
The DE model implemented here assumes perfect within-compartment mixing, implying any infectious individual can meet any susceptible individual with equal probability. Realistic networks are far more sparse and clustered. We explore five different network structures: fully connected, random (Erdos and Renyi 1960), small world (Watts and Strogatz 1998), scale-free (Barabasi and Albert 1999), and lattice (where contact only occurs between neighbors on a ring). We parameterize the model so that all networks (other than the fully connected case) have the same mean number of links per node, $k = 10$ (Watts 1999).

The fully connected network corresponds to the perfect mixing assumption of the DE model. The random network is similar except people are linked with equal probability to a subset of the population. To test the network most different from the perfect mixing case, we also model a one-dimensional ring lattice with no long-range contacts. With $k = 10$, each person

contacts only the five neighbors on each side. The small world and scale-free networks are intermediate cases with many local and some long-distance links. These widely used networks characterize a number of real systems (Watts 1999, Barabasi 2002). We set the probability of long-range links in the small world networks to 0.05, in the range used by Watts (1999). We build the scale-free networks using the preferential attachment algorithm (Barabasi and Albert 1999) in which the probability a new node links to existing nodes is proportional to the number of links each node already has. Preferential attachment yields a power law for the probability that a node has k links, $\text{Prob}(k) = \alpha k^{-\gamma}$. Empirical studies typically show $2 \leq \gamma \leq 3$; the mean value of γ in our experiments is 2.6.

The fully connected and lattice networks are deterministic, so every simulation of these cases has the same network-governing contacts among individuals. The Erdos-Renyi, small world, and scale-free cases are random networks. Each simulation of these cases uses a different realization of the network structure. In realistic networks, the links among individuals change through time even as overall topology can remain stable (e.g., Kossinets and Watts 2006), introducing mixing that brings the AB model closer to the assumptions of the compartment model. To maximize the differences between the AB and DE conditions, however, we assume the network realization in each simulation is fixed. The online supplement details the construction of each network.

Individual Heterogeneity
Each individual has four relevant characteristics: expected contact rate, infectivity, emergence time, and disease duration. In the homogeneous condition ($H_=$), each individual is identical with parameters set to the values of the DE model. In the heterogeneous condition ($H_{\neq}$), we vary individual contact frequencies.

Heterogeneity in contacts is modeled as follows. Given that two people are linked (that they can come into contact), the frequency of contact between them depends on two factors. First, how often does each use their links, on average: some people are gregarious; others shy. Second, time constraints may limit contacts. At one extreme, the frequency of link use

may be constant, so that people with more links have more total contacts per day, a reasonable approximation for some airborne infections and easily communicated ideas: a professor may transmit an airborne virus or a simple concept to many people with a single sneeze or comment, (roughly) independent of class size. At the other extreme, if the time available to contact people is fixed, the chance of using a link is inversely proportional to the number of links, a reasonable assumption when transmission requires extended personal contact: the professor can only tutor a limited number of people each day. We capture these effects by assigning individuals different propensities to use their links, $\lambda[j]$, yielding the expected contact frequency for the link between individuals i and j, $c[i, j]$:

$$c[i, j] = \kappa * \lambda[i] * \lambda[j]/(k[i] * k[j])^{\tau}, \qquad (5)$$

where $k[j]$ is the total number of links individual j has, τ captures the time constraint on contacts, and κ is a constant chosen to ensure that the expected contact frequency for the population equals the value used in the DE model. In the $H_=$ condition, $\lambda[j] = 1$ for all j and $\tau = 1$, so that expected contact frequencies are equal for all individuals, independent of how many links each has. In the $H_{\neq}$ condition, $\lambda[j]$ is a random variable and $\tau = 0$: individuals have different contact rates and those with more links have more contacts per day. We use a uniform distribution, $\lambda[j] \sim U[0.25, 1.75]$.

Calibrating the DE Model
In practice, the parameters determining R_0 are often poorly constrained by biological and clinical data. For emerging diseases such as vCJD, BSE, and avian flu, data are not available until the epidemic has already spread. Parameters are usually estimated by fitting models to aggregate data as an outbreak unfolds; SARS provides a typical example (Dye and Gay 2003, Lipsitch et al. 2003, Riley et al. 2003). Because R_0 also

depends on contact networks that are often poorly known, models of established diseases are commonly estimated the same way (e.g., Gani and Leach 2001). To mimic this protocol, we treat each realization of the AB model as the "real world" and estimate the parameters of the DE to yield the best fit to the cumulative number of cases. We estimate infectivity (i_{ES} and i_{IS}) and incubation time ($1/\varepsilon$) by nonlinear least squares in a large set of individual AB realizations (see the online supplement). Results assess whether calibrated compartment models can capture the behavior of heterogeneous individuals in realistic settings with different contact networks.

Results
For each experimental condition, we examine three measures relevant to public health. The maximum symptomatic infectious population (*peak prevalence*, I_{max}) indicates the peak load on the public health infrastructure including health workers, immunization resources, hospitals, and quarantine facilities. The time from initial exposure to the maximum of the infected population (the *peak time*, T_p) measures how quickly the epidemic spreads and therefore how long officials have to deploy those resources. The fraction of the population ultimately infected (the *final size*, F) measures the total burden of morbidity and mortality. To illustrate, Figure 1 compares the base-case DE model with a typical simulation of the AB model (in the heterogeneous scale-free case). The sample scale-free epidemic grows faster than the DE ($T_p = 37$ versus 48 days), has similar peak prevalence ($I_{max} = 27\%$), and ultimately afflicts fewer people ($F = 85\%$ versus 98%).

In this study, we focus on the practical significance of differences between the mean output of the AB and DE models. Specifically, we explore whether the differences among models are large relative to the variability in outcomes for which policymakers should plan and whether the differences alter the choice of

Figure 1   (Left) DE Model with Base Parameters (Table 1); (Right) Typical Simulation of the Equivalent AB Model with the Heterogeneous Condition of the Scale-Free Network
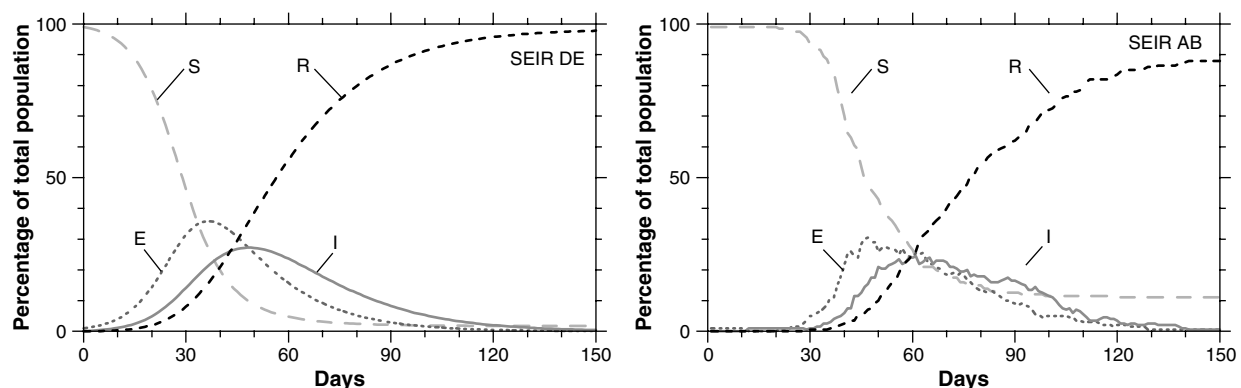
optimal policies. To begin, we conservatively consider outcome variability arising only from stochastic interactions among individuals. Specifically, suppose that policymakers planning for a possible outbreak know with certainty mean infectivity, incubation period, disease duration, network type, and all other parameters conditioning contagion and diffusion, and that these characteristics are unaffected by the course of the epidemic. In short, assume that policymakers possess a perfect agent-based model of the situation, and lack only knowledge of which individuals will, by chance, encounter each other at any moment and transmit the disease. As an example, suppose that the contact network is characterized by a scale-free degree distribution with known parameters, and that individuals are heterogeneous in their behavior (but with known distribution). For the hypothetical disease we examine, prevalence peaks on average after 44 days at a mean of 23.9% of the population. In the deterministic compartment model with the same parameters, prevalence peaks after 48 days at 27.1% of the population. Given the large sample of AB realizations, these differences are statistically significant ($p < 0.001$), but they are not practically significant. Unobservable stochastic interactions among individuals means policymakers, to be, for example, 95% confident resources will be sufficient, must plan to handle an epidemic peaking between 4 and 75 days after introduction, with peak prevalence between 4% and 31.5% of the population. Of course, the deterministic model yields a single trajectory representing the expected path under the mean-field approximation. No responsible policymaker should plan for the mean epidemic without considering uncertainty. To assess the range of outcomes arising from the random nature of individual interactions, policymakers using compartment models would have to estimate the impact of uncertainty by, for example, moving to a stochastic DE representation. Such a model would be computationally efficient relative to the full AB model, but would still assume within-compartment mixing and homogeneous agents.

Policymakers should also consider how model assumptions affect the optimality of interventions. Consider, for example, a quarantine policy. Quarantine should be implemented if its benefit/cost ratio (e.g., the value of QALYs or DALYs saved and avoided health costs relative to the costs of quarantine implementation) is favorable and higher than that of other policy options (including no action). Two models may yield similar estimates of epidemic diffusion, yet respond differently to policies. In such cases, the differences between the models may be of great practical significance even if their base-case behavior is similar. We provide an example below.

Figure 2 shows the symptomatic infectious population, I, in 1,000 AB simulations for each network and heterogeneity condition. Also shown are the mean of the ensemble and the trajectory of the base case DE model. Table 2 reports results for the fitted DE models; Tables 3 and 4 compare the means of T_p, I_{max}, and F for each condition with both the base and fitted DE models. Except for the lattice, the DE and mean AB dynamics are qualitatively similar. Initial diffusion is driven by positive feedback as contagious individuals spread the infection. The epidemic peaks when the susceptible population is sufficiently depleted that the mean number of new cases generated by contagious individuals is less than the rate at which they are removed from the contagious pool.

Departures from the DE model increase from the connected to the random, scale-free, small world, and lattice structures (Figure 2; Tables 3 and 4). The degree of clustering explains some of these variations. In the fully connected and random networks, the chance of contacts in distal regions is the same as for neighbors. The positive contagion feedback is strongest in the connected network because an infectious individual can contact everyone else, minimizing local contact overlap. In contrast, the lattice has maximal clustering. When contacts are localized in a small region of the network, infectious individuals contact their common neighbors repeatedly. As these people become infected, the chance of contacting a susceptible and generating a new case declines, slowing diffusion on average, even if the total susceptible population remains high.

In the deterministic DE model, there is always an epidemic if $R_0 > 1$. Due to the stochastic nature of interactions in the AB model, it is possible that no epidemic occurs or that it ends early if, by chance, the few initially contagious individuals recover before generating new cases. As a measure of early burnout, Table 3 reports the fraction of cases where cumulative cases remain below 10%. (Except for the lattice, the results are not sensitive to the 10% cutoff. The online supplement shows the histogram of final size for each network and heterogeneity condition.) Early burnout ranges from 1.8% in the homogeneous connected case to 6.8% in the heterogeneous scale-free case. Heterogeneity raises the incidence of early burnout in each network because there is a higher chance that the first cases will have few contacts and recover before spreading the disease. Network structure also affects early burnout. Greater contact clustering increases the probability that the epidemic burns out in a local patch of the network before it can jump to other regions, slowing diffusion and increasing the probability of global quenching.

Consider now the differences between the DE and AB cases by network type.

Figure 2 **Prevalence of Symptomatic Infectious Individuals** (I/N, %)
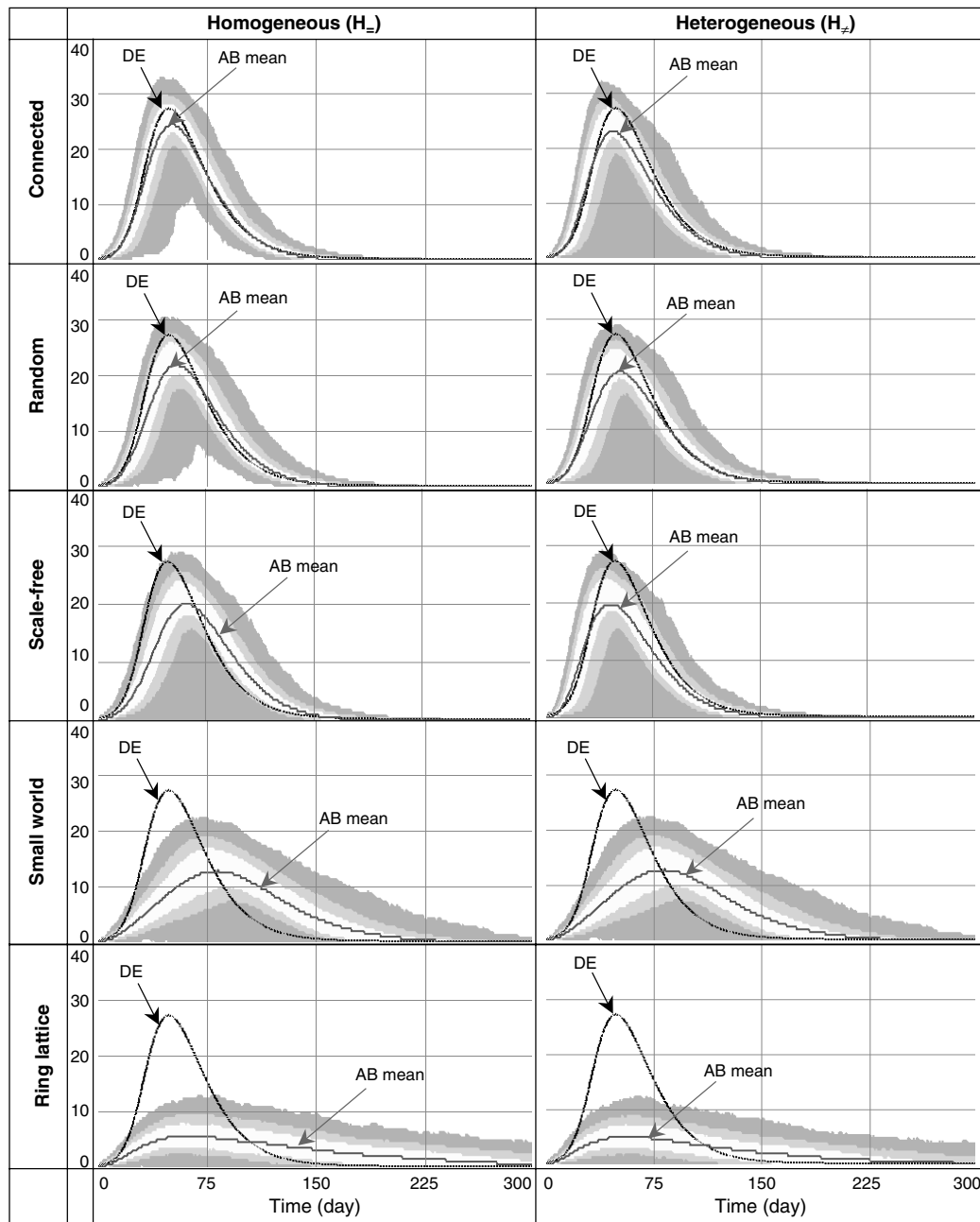
Note. Panels show the envelopes comprising 50%, 75%, and 95% of 1,000 AB simulations for each network and heterogeneity condition, the mean of the AB simulations, and the trajectory of the base-case DE model.

Heterogeneity results in smaller final size, F, in all conditions: the mean reduction over all 10 conditions is 0.10, compared to a mean standard deviation across all conditions, $\bar{\sigma}$, of 0.19. Similarly, heterogeneity reduces T_p in all conditions (by a mean of 9.5 days, with $\bar{\sigma} = 26$ days). Maximum prevalence also falls in all conditions (by 1.5%, $\bar{\sigma} = 5.1\%$). In the $H_{\neq}$ condition, high-contact individuals tend to become infected sooner, causing, on average, faster take-off compared to the $H_=$ case (hence, earlier peak times).

These individuals are also removed sooner, reducing mean contact frequency, and hence the reproduction rate, among those who remain compared to the $H_=$ case. Subsequent diffusion is slower, peak prevalence is smaller, and the epidemic ends sooner, yielding fewer cumulative cases.

Fully Connected Network
The fully connected network corresponds closely to the perfect mixing assumption of the DE. As expected, the base DE model closely tracks the mean of

Table 2 Median Estimated Value of R_0 for the Calibrated DE Models

		Connected		Random		Scale-free		Small world		Lattice	
Parameter		$H_=$	$H_{\neq}$	$H_=$	$H_{\neq}$	$H_=$	$H_{\neq}$	$H_=$	$H_{\neq}$	$H_=$	$H_{\neq}$
Implied $R_0 = c_{ES} * i_{ES}/\varepsilon$	Median	4.21	2.96	3.10	2.54	3.15	2.27	3.35	2.54	1.61	1.35
$+ c_{IS} * i_{IS}/\delta$	σ	1.71	0.62	0.58	0.52	0.66	0.64	0.88	0.66	0.81	0.55
R^2	Median	0.999	0.999	0.999	0.999	0.999	0.999	0.998	0.998	0.985	0.987
	σ	0.025	0.049	0.017	0.050	0.016	0.039	0.040	0.059	0.056	0.043

Note. R^2, the square of the Pearson correlation coefficient, measures goodness of fit between each AB and calibrated DE simulation.

the AB simulations. In the $H_=$ condition, T_p, I_{max}, and F in the base DE model fall well within the 95% confidence interval defined by the ensemble of AB simulations. In the $H_{\neq}$ case, T_p and I_{max} also fall within the 95% range, but F lies just outside the range encompassing 95% of the ensemble.

Random Network
The random network behaves much like the connected case. The DE values of T_p and I_{max} fall within the 95% outcome range for both heterogeneity conditions. The value of F in the DE falls outside the 95% range for both $H_=$ and $H_{\neq}$ because the sparse contact network means more people escape contact

with infectious individuals compared to the perfect mixing case.

Scale-Free Network
The scale-free network departs substantially from perfect mixing. Most nodes have few links, so initial take-off is slower, but once the infection reaches a hub it spreads quickly. The base DE values of T_p and I_{max} fall well within the 95% outcome interval for both heterogeneity conditions. However, as the hubs are removed from the infectious pool, the remaining nodes have lower average contact rates, causing the epidemic to burn out at lower levels of diffusion; the 95% range for final size is 2% to 98% for $H_=$ and 1% to 92% for $H_{\neq}$, while the base DE value is 98%.

Table 3 Mean and Standard Deviation of Final Size, F, in (1) the AB, (2) Fitted DE Simulations, (3) Percentage of AB, and (4) Fitted DE Simulations with $F < 0.10$

	Connected		Random		Scale-free		Small world		Lattice	
	$H_=$	$H_{\neq}$	$H_=$	$H_{\neq}$	$H_=$	$H_{\neq}$	$H_=$	$H_{\neq}$	$H_=$	$H_{\neq}$
(1) AB mean (σ)	0.97 (0.13)	0.90* (0.19)	0.92* (0.15)	0.86** (0.17)	0.92* (0.16)	0.80** (0.22)	0.92 (0.17)	0.83** (0.21)	0.65 (0.29)	0.51** (0.26)
(2) Fitted DE μ (σ)	0.98 (0.07)	0.91 (0.17)	0.92 (0.15)	0.86 (0.18)	0.92 (0.15)	0.78 (0.25)	0.92 (0.17)	0.83 (0.22)	0.62 (0.28)	0.50 (0.26)
(3) AB %$F < 0.10$	1.8	4.4	2.7	3.8	2.9	6.8	2.6	4.8	3.1	5.9
(4) Fitted DE %$F < 0.10$	0.5	3.5	2.5	4.0	2.5	9.0	2.0	5.0	4.0	4.5

*/** indicates that F in the base DE (0.98) falls outside the range encompassing 95%/99% of the AB simulations.

Table 4 Peak Time and Peak Prevalence in 1,000 Simulations of the AB Model and Calibrated DE Models for Each Experimental Condition

	Connected		Random		Scale-free		Small world		Lattice	
	$H_=$	$H_{\neq}$	$H_=$	$H_{\neq}$	$H_=$	$H_{\neq}$	$H_=$	$H_{\neq}$	$H_=$	$H_{\neq}$
Peak time, T_p (days)										
AB μ	49.8	44.9	52.8	49.5	60.6	43.6	86.5	83.6	84.4	75.2
AB σ	10.8	12.4	13.4	14.2	16.7	15.2	31.5	36.4	57	50.9
Fitted DE μ	51.3	49.6	56.6	58	62.9	47.0	82.1	83.2	102.8	90.5
Fitted DE σ	9	21.3	21.3	37.7	23.2	24.6	25.2	34.3	77.8	79.7
Peak prev I_{max} (%)										
AB μ	29.1	27.1	26.5	25.1	24.6	23.9	18.1*	16.5*	9.3**	8.5**
AB σ	4.9	6.3	5.2	5.6	5.2	6.8	4.8	5.3	3.4	3.4
Fitted DE μ	26.9	26.7	24.6	24.2	22.8	21.4	17	14.9	11	7.8
Fitted DE σ	2.8	13.2	9.6	18.1	6.1	9.2	4.4	5	19.9	9.9

*/** indicates the base DE values ($T_p = 48$ days and $I_{max} = 27.1\%$) fall outside the 95%/99% confidence bounds defined by the ensemble of AB simulations.

Small World Network

Small world networks are highly clustered and lack highly connected hubs. Nevertheless, the presence of a few long-range links is sufficient to seed the epidemic throughout the population (Watts and Strogatz 1998). Diffusion is slower on average compared to the DE and the connected, random, and scale-free networks. The existence of a few randomly placed long-range links also increases the variability in outcomes. The 95% range for T_p is 22 to 154 days for H$_=$ (7 to 176 days for H$_{\neq}$), easily encompassing the base DE value. Slower diffusion relative to the DE causes peak prevalence in the DE to fall outside the 95% interval of AB outcomes for both H$_=$ and H$_{\neq}$. The main impact of heterogeneity is greater dispersion and a reduction in final size.

Ring Lattice Network

In the lattice, individuals only contact their k nearest neighbors, so the epidemic advances roughly linearly in a well-defined wave of new cases trailed by symptomatic and then recovered individuals. Such waves are observed in the spread of plant pathogens, where transmission is mostly local, although in two dimensions more complex patterns are common (Bjornstad et al. 2002, Murray 2002). Because the epidemic wave front reaches a stochastic steady state in which removal balances new cases, the probability of burnout is roughly constant over time, and $I_{\max}$ is lower, with the base DE value falling outside the 99% range. For the same reason, mean final size is much lower and peak time is longer than the base DE. Interestingly, the variance is higher as well, so that in the H$_=$ condition, the DE values of F and T_p fall within the 95% range of AB outcomes.

In sum, peak time in the uncalibrated base DE model falls within the envelope encompassing 95% of the AB simulations in all 10 network and heterogeneity conditions. Peak prevalence falls within the 95% range in all but the small world and lattice. Final size, however, is sensitive to clustering and heterogeneity, falling within the 95% range in only three cases.

Calibrated DE Model

In practice, parameters such as R_0 and incubation times are poorly constrained and are estimated by fitting models to aggregate data. Table 2 summarizes the results of fitting the DE model to 200 randomly selected AB simulations in each experimental condition, a total of 2,000 calibrations. The median R^2 for the fit to cumulative cases exceeds 0.985 in all scenarios. The mean values of F, T_p, and $I_{\max}$ in the calibrated DE fall within the range encompassing 95% of the AB outcomes in all network and heterogeneity conditions. The DE model fits well even though it is clearly misspecified in all but the homogeneous fully

connected network. Why? As the network becomes increasingly clustered and diffusion slows, the estimated parameters adjust accordingly. Specifically, in deterministic SEIR compartment models, R_0 and final size are related by $R_0 = -\ln(1 - F)/F$ (Anderson and May 1991). Consequently, when contact clustering leads to smaller F, the estimated incubation time or transmission rates must shift to yield a smaller estimate of R_0. The parameter estimates are biased because deviations from their underlying values are the only way the DE, with its within-compartment homogeneity and mixing assumptions, can capture the impact of heterogeneity and network structure. Further, the close fit of the compartment model does not imply that its response to policies will be the same as that of the underlying clustered and heterogeneous network. The online supplement provides further details.

Sensitivity to Population Size

We repeated the analysis for $N = 50$ and 800 (see the online supplement). The results change little over this factor of 16. For most conditions, the rate of early burnout falls in the larger population, so the final fraction of the population infected is slightly larger (and therefore closer to the value in the DE). Population size has little impact on the other metrics.

Sensitivity to R_0

We varied R_0 from 0.5 to 2 times the base value; detailed results are reported in the online supplement. Naturally, diffusion is strongly affected by R_0. Somewhat surprisingly, however, over the range tested the differences between the DE and mean AB outcomes remain small relative to the 95% outcome range for most of the metrics. Changes in R_0 have two offsetting effects. First, the smaller the value of R_0, the larger are the differences between the DE and means of the AB trajectories. Second, however, the smaller R_0, the greater the variation in outcomes within a given network and heterogeneity condition caused by chance encounters among individuals. Small values of R_0 reduce the expected number of new cases each infectious individual generates before removal. In effect, the fraction of the contact network sampled by each infectious individual is smaller, so the probability that the epidemic will be seeded at multiple points in the network decreases. In highly clustered and heterogeneous networks, the lower representativeness of these small samples increases the difference between the DE and the mean of the AB trajectories (for example, more cases of early quenching will be observed). For the same reason, however, individual realizations of the same network and heterogeneity condition will differ more with small R_0, increasing the variance in outcomes for which policymakers must

prepare. Similarly, larger R_0 reduces the differences between the DE model and the means of the AB models, but also reduces variability in outcomes because each infectious individual samples the network many times before recovering. These offsetting effects imply that, over the range examined here, the differences between the DE and the mean behavior of the AB models are relatively insensitive to variations in R_0.

Sensitivity to Disease Natural History
In many diseases, the exposed gradually become more infectious prior to becoming symptomatic. This progression can be modeled by adding additional compartments to the exposed stage with different infectivities in each. In the classic SEIR model used here, with only one compartment per stage, presymptomatic infectivity is approximated by assuming the exposed are contagious, though with $i_{ES} < i_{IS}$. To test the impact of this assumption, we set $i_{ES} = 0$, adjusting i_{IS} to keep R_0 at its base value. Results are reported in the online supplement. As expected, diffusion slows and the probability of early quenching grows. However, the differences in the mean values of the metrics across models generally remain small relative to the 95% range of AB outcomes. Assuming that exposed individuals are not contagious has little impact on the differences between the DE and mean AB behavior relative to the variability in AB outcomes.

Policy Analysis and Sensitivity to Model Boundary
Another important question is whether the behavior of the models differs in response to policy interventions and expansion of the model boundary. While comprehensive treatment of these questions is beyond the scope of this paper, we illustrate by examining the impact of actions that reduce contact rates. For example, the 2003 SARS epidemic appears to have been quenched through contact reduction (Wallinga and Teunis 2004, Riley et al. 2003, Lipsitch et al. 2003). Contact reduction can arise from policies, e.g., quarantine (including mandatory isolation and travel restrictions), and from behavioral feedbacks, e.g., social distancing, where individuals who fear infection reduce contacts with others. For simplicity, we assume that contact rates fall linearly to a minimum value as the total number of confirmed cases (cumulative prevalence $P = I + R$) rises.[2] Specifically, we model the contact frequency c_{js} between infectious persons, $j \in \{E, I\}$, and susceptibles, $s \in \{S\}$, as a

weighted average of the initial frequency, c_{js}^*, and the minimum achieved under quarantine, c_{js}^q:

$$c_{js} = (1 - q)c_{js}^* + qc_{js}^q, \qquad (6)$$

$$q = \text{Min}[1, \text{Max}(0, (P - P_0)/(P_q - P_0))]. \qquad (7)$$

The impact of contact reduction, q, rises linearly as cumulative prevalence, P, rises from a threshold, P_0, to the level at which the effect saturates, P_q. We set $P_0 = 2$ and $P_q = 10$ cases. Neither social distancing nor quarantine are perfect; we set the minimum contact frequency, $c_{js}^q = 0.15c_{js}^*$. This value gradually reduces R_0 in the DE model from 4.125 to ≈ 0.6, roughly similar to the reduction Wallinga and Teunis (2004) estimate for the SARS epidemic.

As expected, contact reduction quenches the epidemic earlier. In the DE model, prevalence peaks 17 days sooner, I_{max} falls from 27% to 4.4%, and F falls from 98% to 19% of the population, greatly easing the burden on public health resources. Contact reduction has similar benefits in the AB cases. The differences between the means of the metrics in the AB models and their DE value are small relative to the variation in outcomes caused by stochastic interactions in the AB models. The DE results fall within the 95% outcome range for all three metrics in all network and heterogeneity conditions, with one exception: the value of F in the lattice (Table 5). However, as in the base case, clustering and heterogeneity cause some differences between the DE and mean AB outcomes. Under contact reduction, heterogeneity increases mean F because high-contact individuals tend to be infected first, increasing the exposed population relative to $H_=$ before contact reduction is triggered. In the base case, however, heterogeneity lowers F because early high-contact cases are also removed early, lowering the reproduction rate. Therefore, the mean reduction in F under contact reduction is smaller in the heterogeneous cases.

Policies should be implemented if their cost-benefit ratio is favorable compared to other options, including no action. As a simple illustration, suppose that the per-capita costs of mandatory contact reduction policies, denoted C, are fixed and that the benefits are linear in avoided cases, $\Delta F = F_{\text{no policy}} - F_{\text{policy}}$. Ignoring uncertainty, and hence issues of policymaker risk aversion, mandatory measures should be implemented if $b\Delta F > C$, where b is the benefit per avoided case. In the scale-free case, $\Delta F = 0.75$ for the $H_=$ case, but falls to 0.59 in the $H_{\neq}$ condition (see the online supplement). For $0.59 \leq C/b \leq 0.75$, whether mandatory measures are indicated depends on whether the population is homogeneous or not. Uncertainty, nonlinear costs and benefits, or risk-averse policymakers will change the width of this interval of policy sensitivity, but not the principle that the choice among

---

[2] Other policies, such as targeted immunization, can exploit the structure of the contact network, if it is known, and generally require an AB model, although some such policies can be approximated in DE models (e.g., Kaplan et al. 2003).

Table 5   Public Health Metrics Under Contact Reduction

Metric		Connected H$_=$	Connected H$_{\neq}$	Random H$_=$	Random H$_{\neq}$	Scale-free H$_=$	Scale-free H$_{\neq}$	Small world H$_=$	Small world H$_{\neq}$	Lattice H$_=$	Lattice H$_{\neq}$
Final size F	μ	0.215	0.249	0.157	0.201	0.148	0.247	0.112	0.117	0.102*	0.099*
	σ	0.084	0.091	0.064	0.088	0.062	0.105	0.044	0.048	0.037	0.035
Peak time T_p	μ	35.0	36.1	33.1	34.6	34.1	34.9	30.3	30.5	29.4	30.4
	σ	15.3	15.9	14.8	16.7	15.7	18.1	13.5	14.2	13.5	14.6
Peak prev I_{max}	μ	6.42	7.28	5.17	6.15	4.98	7.43	4.17	4.35	3.97	3.89
	σ	2.40	2.66	1.99	2.55	1.98	3.23	1.58	1.67	1.52	1.42

*/** indicates the DE simulation falls outside the 95%/99% confidence bound defined by the ensemble of AB simulations. The results for the DE model are $F = 0.190$, $T_p = 31.3$ days, and $I_{max} = 4.43\%$.

policies may be sensitive to network type, individual heterogeneity, and other assumptions.

The size of the region of policy sensitivity also depends on the model boundary. For example, if awareness of the epidemic arising from, e.g., media reports causes individuals to engage in social distancing spontaneously, contacts will fall even without quarantine and travel restrictions, reducing the benefits of mandatory measures. If spontaneous social distancing reduces R_0 persistently below one, mandatory measures would not be needed to quench the epidemic and would not be justified on cost-benefit grounds. At the other extreme, if the public's reaction to media reports was panic and flight, increasing the risk of long-range transmission, mandatory control measures would have even higher benefits relative to their costs. Thus, policymakers should carry out sensitivity analysis not only over uncertainty in parameters, network topology, and individual characteristics, but over variations in the strength of behavioral feedback effects, that is, over a wide model boundary. Additional resources for empirical work to reduce uncertainty and to improve the model should be allocated where they have the greatest value. Making such judgments rationally requires the resources and time to carry out sensitivity analysis for each policy option across all relevant dimensions of uncertainty.

Discussion and Conclusions

Stimulated by advances in computation, agent-based simulation is growing in popularity. Still, no matter how powerful computers become, limited time, budget, cognitive capabilities, and decision-maker attention mean modelers always face trade-offs: should they disaggregate to capture the diverse attributes of individuals, expand the model boundary to capture additional feedback processes, or keep the model simple so that it can be analyzed thoroughly? Simple models can be analyzed thoroughly, but may lack the structure needed to observe important

dynamics and fully inform policymakers. For example, compartment models are computationally efficient, but assume perfect mixing and homogeneity within compartments. Agent-based models increase computational requirements, constraining sensitivity analysis, but easily capture the networks of relationships among individuals and heterogeneity in their attributes. By contrasting agent-based and compartment models of epidemic diffusion, we assess the importance of relaxing the perfect mixing and homogeneity assumptions.

As expected, network topology and individual heterogeneity affect the dynamics. Higher clustering increases the overlap in contacts among neighbors and therefore slows diffusion to other regions, leading, on average, to lower peak prevalence and higher peak times in the small world and lattice networks. Heterogeneity in individual contact rates causes slightly earlier mean peak times as high-contact individuals rapidly seed the epidemic, followed by lower diffusion levels as the high-contact individuals are removed, leaving those with lower average transmission probability and a smaller reproduction rate. These results are consistent with analysis of heterogeneity for SI and SIS models (Veliov 2005). Such dynamics were also observed in the HIV epidemic, where initial diffusion was rapid in subpopulations with high contact rates. Finally, in the stochastic AB models, the epidemic fizzles out in a small fraction of cases even though the underlying parameters yield an expected value for the basic reproduction rate greater than one. The more highly clustered and heterogeneous the population, the greater is the incidence of early quenching. The deterministic DE model cannot generate such behavior.

Before turning to implications, we consider limitations and extensions. The experiments examined the classic SEIR model. Further work should address the robustness of results to common elaborations such as loss of immunity, nonexponential distributions for emergence and recovery, recruitment of new susceptibles, nonhuman disease reservoirs and vectors, etc.

Note, however, that by using the deterministic SEIR model, with only four compartments, we maximize the difference between the aggregation assumptions of the DE and AB representations. In practical applications, DE models often disaggregate the population more finely to account for clustering, heterogeneity, and other attributes that vary across individuals (e.g., age, sex, location). AB models capture heterogeneity directly by assigning attributes to individuals. Adding compartments to DE models allows modelers to approximate the heterogeneity and clustering in a situation, while retaining the computational advantages of the compartmental paradigm (Riley 2007). A challenge for future work is optimally choosing the number and definitions of compartments to capture the impact of clustering and heterogeneity. Comparing disaggregated DE models to AB models may be useful in designing stochastic compartment models that capture heterogeneity and network effects using the fewest additional compartments, if the network structure is known and stable.[3] Testing this proposal is beyond the scope of this paper.

Although we examined a wide range of networks, the AB models contain many parameters that could be subject to additional sensitivity analysis, including the mean number of links per node, the probability of long-range links (in the small world network), and the scaling exponent (in the scale-free case). Other dimensions of heterogeneity and other networks could be examined, including networks derived from field study (Ahuja and Carley 1999). The number and distribution of the initially infectious individuals can be varied. The robustness of other policies with respect to network type, heterogeneity, and model boundary should be examined. The boundary could be expanded to include endogenously the many effects that alter contact rates and network structure as an epidemic progresses (relaxing the fixed-network assumption).

The deterministic DE model does not capture the variability in outcomes caused by stochastic events and yields a point value for any metric (for a given set of parameters). The costs and benefits of options facing policymakers, however, often depend on the distribution of possible outcomes, not only expected values. It is not appropriate to use the output from a single run of a deterministic model to answer policy questions such as, "Is hospital capacity sufficient to handle an outbreak?" Stochastic compartment models, along with individual-level AB models, can capture the distribution of outcomes generated by random interactions among individuals and should be tested against the full AB models. Indeed, given the many sources of uncertainty in realistic settings, it

would be irresponsible to use a single set of assumptions in any model, deterministic or stochastic, compartment or agent-based. The outcome distributions in the AB results reported here only capture uncertainty arising from stochastic events; sensitivity to parameters may be larger and should be examined. For example, R_0 for smallpox is estimated to be in the range 3–6 (Gani and Leach 2001). Varying R_0 over that range in the DE (by scaling both contact rates, c_{ES} and c_{IS}, proportionately) causes F to vary from 94.1% to 99.7%, $I_{\max}$ to vary from 21.7% to 31.8%, and T_p to vary from 36 to 64 days, comparable to the differences caused by relaxing the perfect mixing and homogeneity assumptions of the compartment model. Such uncertainty further highlights the importance of wide-ranging parametric and structural sensitivity tests for all models.

Finally, the results may inform phenomena beyond epidemics. Processes of social contagion (imitation, word of mouth, etc.) play important roles in many social and economic phenomena, from marketing to crowd behavior (Strang and Soule 1998, Rogers 2003). Models of diffusion in such contexts are similar to the SEIR family, most notably the Bass (1969) model and its extensions (e.g., Mahajan et al. 2000). Moreover, modelers tackling policy issues related to innovation and product diffusion face trade-offs in the choice of modeling assumptions similar to those studying epidemics (Gibbons 2004). We do not explicitly address the differences between AB and DE models in these contexts; such issues are an important arena for future work.

The results demonstrate a number of differences between the deterministic compartment model and the individual-level models, and across the AB models with different network and heterogeneity assumptions. The significance of these differences depends on the purpose of the model and specifics of the situation. Here we focus on policy analysis where resources are limited and policymakers must make trade-offs among the level of detail, the breadth of the model boundary, and the ability to carry out sensitivity analysis. As expected, the differences between the mean behavior of the stochastic agent-based models and the deterministic compartment model are statistically significant when the homogeneity and perfect mixing assumptions of the compartment model are violated. However, these differences may have little practical significance in some settings. In the cases tested here, the differences in the peak burden on public health resources and the time available to deploy those resources between the DE and the mean of the AB models are small relative to the variability in outcomes caused by unobservable stochastic interactions among individuals for the connected, random, small world, and scale-free networks. The DE values

---

[3] We thank an anonymous reviewer for this suggestion.

of these metrics generally fall within the envelope capturing 95% of the AB realizations, not only in the base case but across a range of assumptions about R_0, population size, and disease life cycle. The main exception is the ring lattice, where there are no long-range contacts. However, a pure lattice is unrealistic in modeling human diseases due to the high mobility of modern society (although it may be appropriate in modeling immobile plant populations or where transmission to humans arises only from geographically constrained vectors).

Because parameters characterizing emerging (and some established) diseases are poorly constrained, epidemiologists typically fit models to aggregate data for particular outbreaks. We tested the impact of this protocol by treating each realization of the AB model as the "real world" and fitting the DE model to them. The calibrated compartment model captures the dynamics well, with the median R^2 exceeding 0.985 in all conditions. The means of the public health metrics for the calibrated models fell within the 95% confidence range defined by the ensemble of AB simulations in all network and heterogeneity conditions tested. These results suggest that simple DE models can capture a wide range of variation in network structure and individual attributes. However, the parameter values obtained by fitting the aggregate model to the data from an AB simulation (and therefore from the real world) do not necessarily equal the mean of the individual-level parameters governing the interactions among individuals. Aggregate parameter estimates not only capture the mean of individual attributes such as contact rates, but also the impact of heterogeneity and network structure. Unless compartment models are sufficiently disaggregated to capture the individual heterogeneity and clustering in the actual contact network, the compartment model will be misspecified and the parameter estimates biased (See Table 2 and the online supplement Table EC.4). Modelers often use both micro and aggregate data to parameterize both AB and DE models. The estimation results suggest that caution must be exercised in doing so, and in comparing parameter values across different models (Fahse et al. 1998). Further, as shown in the online supplement, the ability to reproduce historical data does not imply that calibrated compartment models will respond to policies the same way the AB models do.

AB models enable analysts to examine questions not easily modeled in the DE paradigm, e.g., creating and removing nodes and links to simulate random failures or targeted attacks. AB models can show how aggregate behavior emerges from interactions among the elements of the system (e.g., Reynolds 1987), allow for more realistic representation and analysis of

stochastic behavior in a population, and extend theoretical understanding by identifying instances where DE representations cannot generate certain behaviors (e.g., Shnerb et al. 2000).

Data availability significantly affects model choice. AB models will be useful when data or the underlying "physics" of a situation specify the network structure, suggest it is critical in the results, and that structure is stable over the time horizon of interest. Often, though, data on contact networks and the distribution of individual attributes are hard to obtain and highly uncertain, requiring extensive sensitivity analysis to ensure robust results.

In this study, we focus on the practical significance of differences among models. We compare differences in the mean values of important public health metrics in the different models relative to the uncertainty in outcomes for which policymakers should prepare. Another important measure of practical significance is the impact of model type on the determination of optimal policy. Two models may generate similar base-case behavior, yet respond differently to policies; cost-benefit analysis may therefore lead to different policy choices in different models. For example, contact reduction, whether resulting from mandatory quarantine or voluntary social distancing, reduces the size and impact of the outbreak. However, the results show that network type and individual heterogeneity affect the benefits of contact reduction. Thus, for certain values of the costs and benefits, the decision to implement mandatory control measures such as quarantine will depend on the network structure of and contact heterogeneity within the population. Similarly, the costs and benefits of policies will depend on behavioral feedbacks such as the degree of endogenous social distancing. Sensitivity analysis over these and other sources of uncertainty is required to assess the robustness of policy choices to model assumptions.

Model complexity can be expanded in different directions. Modelers can add detail, disaggregating populations by location, individual attributes, and relationship networks. Alternatively, they can expand the model boundary to include feedbacks with other subsystems. For example, the results reported here assume fixed network structure, contact rates, and infectivities. All are actually endogenous. As prevalence increases, people change their behavior. Social distancing and safer practices disrupt contact networks, reduce contact frequencies, and cut the probability of transmission. From staying home, increased hand washing, and use of masks (for SARS) to abstinence, condom use, and needle exchange (for HIV), endogenous behavior change lowers R_0 and can have large effects on disease diffusion (Blower et al. 2000). Alternatively, behavior change may worsen

an epidemic: people fleeing a disease make contact tracing more difficult and may seed outbreaks in remote areas; more effective treatments for HIV/AIDS increase risky behaviors for some people (Lightfoot et al. 2005). The impact of such feedback effects may be larger than the impact of network structure and individual heterogeneity and should not be excluded in favor of greater detail. The contact reduction test above illustrates. The drop in R_0 with cumulative cases can be interpreted as endogenous social distancing. This feedback has a large impact on mean outcomes compared to the differences between the DE and mean AB outcomes. Expanding the boundary of a model can have effects much greater than those introduced by disaggregation from compartments to individuals.

In a review entitled "Uses and abuses of mathematics in biology," May (2004, p. 793) calls for balance in model development:

> Perhaps most common among abuses, and not always easy to recognize, are situations where mathematical models are constructed with an excruciating abundance of detail in some aspects, whilst other important facets of the problem are misty or a vital parameter is uncertain to within, at best, an order of magnitude. It makes no sense to convey a beguiling sense of "reality" with irrelevant detail, when other equally important factors can only be guessed at.

While further work is needed, the results reported here may be useful to modelers seeking the appropriate balance among detail, scope, and the ability to carry out sensitivity analysis over the inevitable uncertainties we all face.

Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://mansci.journal.informs.org/.

References

Abbey, H. 1952. An examination of the reed-frost theory of epidemics. *Human Biol.* **24**(3) 201–233.

Ahuja, M., K. Carley. 1999. Network structure in virtual organizations. *Organ. Sci.* **10**(6) 741–757.

Anderson, R. M., R. M. May. 1991. *Infectious Diseases of Humans Dynamics and Control*. Oxford University Press, Oxford, UK.

Axelrod, R. 1997. The dissemination of culture—A model with local convergence and global polarization. *J. Conflict Resolution* **41**(2) 203–226.

Axtell, R., J. Axelrod, J. M. Epstein, M. Cohen. 1996. Aligning simulation models: A case study and results. *Computational Math. Organ. Theory* **1**(2) 123–141.

Axtell, R., J. Epstein, J. S. Dean, G. J. Gumerman, A. C. Swedlund, J. Harburger, S. Chakravarty, R. Hammond, J. Parker, M. Parker. 2002. Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley. *Proc. Natl. Acad. Sci. USA* **99**(Suppl. 3) 7275–7279.

Ball, F., D. Mollison, G. Scalia-Tomba. 1997. Epidemics with two levels of mixing. *Ann. Appl. Probab.* **7**(1) 46–89.

Barabasi, A. L. 2002. *Linked: How Everything Is Connected to Everything Else and What It Means*. Perseus, Cambridge, MA.

Barabasi, A. L., R. Albert. 1999. Emergence of scaling in random networks. *Science* **286**(5439) 509–512.

Bass, F. 1969. A new product growth model for consumer durables. *Management Sci.* **15**(5) 215–227.

Bjornstad, O., M. Peltonen, A. M. Liebhold, W. Baltensweiler. 2002. Waves of larch budmoth outbreaks in the European Alps. *Science* **298**(5595) 1020–1023.

Blower, S. M., H. B. Gershengorn, R. M. Grant. 2000. A tale of two futures: HIV and antiretroviral therapy in San Francisco. *Science* **287**(5453) 650–654.

Carley, K. 1992. Organizational learning and personnel turnover. *Organ. Sci.* **3**(1) 20–46.

Chen, L. C., B. Kaminsky, T. Tummino, K. M. Careley, E. Casman, D. Fridsma, A. Yahja. 2004. Aligning simulation models of smallpox outbreaks. *Proc. 2nd Sympos. on Intelligence and Security Informatics, Tucson, AZ*, 1–16.

Davis, G. 1991. Agents without principles? The spread of the poison pill through the intercorporate network. *Admin. Sci. Quart.* **36** 583–613.

Dye, C., N. Gay. 2003. Modeling the SARS epidemic. *Science* **300**(5627) 1884–1885.

Edwards, M., S. Huet, F. Goreaud, G. Deffuant. 2003. Comparing an individual-based model of behavior diffusion with its mean field aggregate approximation. *J. Artificial Societies Soc. Simulation* **6**(4). http://jasss.soc.surrey.ac.uk/6/4/9.html.

Epstein, J. M. 2006. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton University Press, Princeton, NJ.

Epstein, J. M., D. Cummings, S. Chakravarty, R. M. Singa, D. S. Burke. 2004. *Toward a Containment Strategy for Smallpox Bioterror—An Individual-Based Computational Approach*. Brookings Institution Press, Washington, D.C.

Erdos, P., A. Renyi. 1960. On the evolution of random graphs. *Publ. Math. Inst. Hungarian Acad. Sci.* **5** 17–61.

Eubank, S., H. Guclu, V. S. A. Kumar, M. V. Marathe, A. Srinivasan, Z. Toroczkai, N. Wang. 2004. Modelling disease outbreaks in realistic urban social networks. *Nature* **429**(6988) 180–184.

Fahse, L., C. Wissel, V. Grimm. 1998. Reconciling classical and individual-based approaches in theoretical population ecology: A protocol for extracting population parameters from individual-based models. *Amer. Naturalist* **152**(6) 838–852.

Ferguson, N. M., M. J. Keeling, J. W. Edmunds, R. Gani, B. T. Grenfell, R. M. Anderson, S. Leach. 2003. Planning for smallpox outbreaks. *Nature* **425**(6959) 681–685.

Gani, J., S. Yakowitz. 1995. Error bounds for deterministic approximations to Markov processes, with applications to epidemic models. *J. Appl. Probab.* **32**(4) 1063–1076.

Gani, R., S. Leach. 2001. Transmission potential of smallpox in contemporary populations. *Nature* **414**(6865) 748–751.

Gibbons, D. E. 2004. Network structure and innovation: Ambiguity effects on diffusion in dynamic organizational fields. *Acad. Management J.* **47**(6) 938–951.

Glass, T., M. Schoch-Spana. 2002. Bioterrorism and the people: How to vaccinate a city against panic. *Clinical Infectious Diseases* **34**(2) 217–223.

Greenhalgh, D., F. Lewis. 2001. Stochastic models for the spread of HIV amongst intravenous drug users. *Stochastic Models* **17**(4) 491–512.

Halloran, E., I. Longini, N. Azhar, Y. Yang. 2002. Containing bioterrorist smallpox. *Science* **298**(5597) 1428–1432.

Hethcote, H. W. 2000. The mathematics of infectious diseases. *SIAM Rev.* **42**(4) 599–653.

Jacquez, J. A., P. O'Neill. 1991. Reproduction numbers and thresholds in stochastic epidemic models. I. Homogeneous populations. *Math. Biosciences* **107**(2) 161–186.

Jacquez, J. A., C. P. Simon. 1993. The stochastic SI model with recruitment and deaths.1. Comparison with the closed SIS model. *Math. Biosciences* **117**(1–2) 77–125.

Jacquez, J. A., C. P. Simon. 2002. Qualitative theory of compartmental systems with lags. *Math. Biosciences* **180** 329–362.

Kaplan, E. H., L. M. Wein. 2003. Smallpox bioterror response. *Science* **300**(5625) 1503.

Kaplan, E. H., D. Craft, L. Wein. 2002. Emergency response to a smallpox attack: The case for mass vaccination. *Proc. Natl. Acad. Sci. USA* **99**(16) 10935–10940.

Kaplan, E. H., D. Craft, L. Wein. 2003. Analyzing bioterror response logistics: The case of smallpox. *Math. Biosciences* **185**(1) 33–72.

Kasperson, R., O. Renn, P. Slovic, H. S. Brown, J. Emel, R. Globle, J. X. Kasperson, S. Ratick. 1988. The social amplification of risk: A conceptual framework. *Risk Anal.* **8**(2) 177–187.

Keeling, M. 1999. The effects of local spatial structure on epidemiological invasions. *Proc. Roy. Soc. London Series B-Biological Sci.* **266**(1421) 859–867.

Koopman, J. S. 2002. Controlling smallpox. *Science* **298**(5597) 1342–1344.

Koopman, J. S., G. Jacquez, S. E. Chick. 2001. New data and tools for integrating discrete and continuous population modeling strategies. *Ann. New York Acad. Sci.* **954**(1) 268–294.

Koopman, J. S., S. E. Chick, C. P. Simon, C. S. Riolo, G. Jacquez. 2002. Stochastic effects on endemic infection levels of disseminating versus local contacts. *Math. Biosciences* **180**(1–2) 49–71.

Kossinets, G., D. Watts. 2006. Empirical analysis of an evolving social network. *Science* **311**(5757) 88–90.

Levinthal, D., J. G. March. 1981. A model of adaptive organizational search. *J. Econom. Behav. Organ.* **2**(4) 307–333.

Lightfoot, M., D. Swendeman, M. J. Rotheram-Borus, W. S. Comulada, R. Weiss. 2005. Risk behaviors of youth living with HIV pre- and post-HAART. *Amer. J. Health Behav.* **29**(2) 162–171.

Lipsitch, M., T. Cohen, B. Cooper, J. M. Robins, S. Ma, L. James, G. Gopalakrishna, S. K. Chew, C. C. Tan, M. H. Samore, D. Fisman, M. Murray. 2003. Transmission dynamics and control of severe acute respiratory syndrome. *Science* **300**(5627) 1966–1970.

Lomi, A., E. R. Larsen. 2001. *Dynamics of Organizations Computational Modeling and Organization Theories*. AAAI Press/MIT Press, Menlo Park, CA.

Mahajan, V., E. Muller, Y. Wind. 2000. *New-Product Diffusion Models*. Kluwer Academic Publishers, Boston.

May, R. M. 2004. Uses and abuses of mathematics in biology. *Science* **303**(5659) 790–793.

McCloskey, D., S. Ziliak. 1996. The standard error of regressions. *J. Econom. Literature* **34**(1) 97–114.

Meadows, D. H., J. Robinson. 1985. *The Electronic Oracle Computer Models and Social Decisions*. John Wiley & Sons, Chichester, UK.

Murray, J. D. 2002. *Mathematical Biology*. Springer, New York.

Newman, M. 2002. Spread of epidemic disease on networks. *Physical Rev. E* **66**(1) 016128-1–016128-11.

Newman, M. 2003. The structure and function of complex networks. *SIAM Rev.* **45**(2) 167–256.

Reynolds, C. 1987. Flocks, herds, and schools: A distributed behavioral model. *Comput. Graphics* **21**(4) 25–34.

Riley, S. 2007. Large-scale spatial-transmission models of infectious disease. *Science* **316**(5627) 1298–1301.

Riley, S., C. Fraser, C. A. Donnelly, A. C. Ghani, L. J. Abu-Raddad, A. J. Hedley, G. M. Leung, L. M. Ho, T. H. Lam, T. Q. Thach, P. Chau, K. P. Chan, P. Y. Leung, T. Tsang, W. Ho, K. H. Lee, E. M. C. Lau, N. M. Ferguson, R. M. Anderson. 2003. Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of public health interventions. *Science* **300**(5627) 1961–1966.

Rogers, E. 2003. *Diffusion of Innovations*. Free Press, New York.

Schelling, T. 1978. *Micromotives and Macrobehavior*. Norton, New York.

Shnerb, N., Y. Louzoun, E. Bettelheim, S. Solomon. 2000. The importance of being discrete: Life always wins on the surface. *Proc. Natl. Acad. Sci. USA* **97**(19) 10322–10324.

Sterman, J. 2000. *Business Dynamics Systems Thinking and Modeling for a Complex World*. Irwin/McGraw-Hill, New York.

Strang, D., S. Soule. 1998. Diffusion in organizations and social movements from hybrid corn to poison pills. *Annual Rev. Sociol.* **24** 265–290.

Tesfatsion, L. 2002. Economic agents and markets as emergent phenomena. *Proc. Natl. Acad. Sci. USA* **99**(Suppl. 3) 7191–7192.

Urban, G., J. Hauser, J. Roberts. 1990. Prelaunch forecasting of new automobiles. *Management Sci.* **36**(4) 401–421.

Veliov, V. 2005. On the effect of population heterogeneity on dynamics of epidemic diseases. *J. Math. Biol.* **51**(2) 123–143.

Wallinga, J., P. Teunis. 2004. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *Amer. J. Epidemiology* **160**(6) 509–516.

Watts, D. 1999. *Small Worlds*. Princeton University Press, Princeton, NJ.

Watts, D. 2004. The new science of networks. *Annual Rev. Sociol.* **30** 243–270.

Watts, D., S. Strogatz. 1998. Collective dynamics of small-world networks. *Nature* **393**(6684) 440–442.