

DETERMINISTIC AND STOCHASTIC MODELS OF INFECTIOUS DISEASE: CIRCULAR MIGRATIONS AND HIV TRANSMISSION DYNAMICS

NATHAN GILL

ABSTRACT. We present two different types of models, deterministic compartmental (based on ordinary differential equations) and stochastic network (based on random graphs), used in the field of population-level HIV transmission modeling, and use them to investigate the impact of circular migration (periodic migrations between urban and rural segments of a larger population) frequency on HIV prevalence. We find that both classes of model, given our particular specifications, predict that circular migration frequency has no impact on HIV prevalence.

CONTENTS

1. Introduction	1
2. Deterministic Models	2
SIR Models	2
The Circular Migrations Model	3
Results	4
3. Network Models	5
Exponential Random Graphs	6
Circular Migration Network Model	7
Results	8
Acknowledgments	9
References	9

1. INTRODUCTION

In the study of population-level HIV transmission dynamics, the utility of modeling is twofold. In an empirical study of a real-world transmission network (e.g. that of gay men on the south side of Chicago), fitting a computer model to experimental data has the potential to reveal patterns in the network that may not have been apparent from the raw data alone. Such data, however, are often difficult to collect. Surveying a large population takes time, and the resulting reconstructed network will almost always be incomplete, since a complete network would require both participation and total honesty from every member of the population in question.

Theoretical modeling can be used to study HIV transmission dynamics in populations that are difficult or impossible to survey. While a theoretical model can never truly capture the intricacies of a living, breathing, social network, one can hope that with thoughtful model design, one can at least capture enough major features of the network to make qualitative statements about the dynamics of infection. As the statistician George Box famously remarked, "all models are wrong, but some are useful"; our model of a particular population may not be perfect, but if it predicts, for example, that a certain kind of public health intervention has a much larger effect on disease prevalence than another, it can still be useful to policy-makers.

In this paper, we use theoretical modeling to study the transmission dynamics of a population split between two separate locations. Certain members of the population migrate between these two locations at regular intervals in a "circular migration" pattern. The model is inspired by migrant populations in sub-Saharan Africa traveling between urban and rural populations, although it will become clear when we

consider the models in detail that the underlying structure could be used to model any split population with migrations of this kind, provided inputs and parameters are chosen properly.

We use two types of models to investigate the impact of migration rate on the overall disease prevalence in our population. The first, deterministic compartmental models, are built around a set of differential equations. The second, stochastic network models, are built around random graphs. In the sections below, we first explain the general theory and principles behind each class of model, and then discuss the details of the corresponding circular migrations model.

2. DETERMINISTIC MODELS

The first class of model we will examine is the deterministic compartmental model. The word “deterministic” signifies that the predictions of these models are determined entirely by their initial conditions, the set of underlying equations, and the input parameter values. In our case, we model the spread of disease using a system of ordinary differential equations whose trajectories depend on parameters like initial number infected, probability of infection per sexual act, average duration of a sexual partnership, etc. Before we develop the relatively more involved equations governing the transmission of HIV in the circular migrations context, we will begin with a very general and basic epidemiological model in order to point out key features in a simple context.

SIR Models. One of the simplest models of the spread of infectious disease is the Susceptible, Infected, Recovered (sometimes Removed), or “SIR,” model. In it, we have three “compartments” for the three different disease status classes, denoted S , I , and R , and the equations relate to the flows among these compartments. The equations are:

$$\begin{aligned}\frac{dS}{dt} &= \frac{r \log(1-c)SI}{N} \\ \frac{dI}{dt} &= \frac{-r \log(1-c)SI}{N} - \gamma I \\ \frac{dR}{dt} &= \gamma I,\end{aligned}$$

where r is the rate of (disease-transmitting) contact among individuals in the population, c is the probability of disease transmission per contact, γ is the rate of recovery from infection, and $N = S + I + R$ is the total size of the population. To see where these equations come from, consider a particular susceptible individual in the population. In a short time interval dt , he makes $r \frac{I}{N} dt$ contacts with infected members of the population, and in each such contact, he has probability $1 - c$ of avoiding the contracting the disease. Thus, the probability that he does not get the disease in the time interval dt is

$$(1 - c)^{r \frac{I}{N} dt},$$

so the probability dp of disease contraction is

$$\begin{aligned}dp &= 1 - (1 - c)^{r \frac{I}{N} dt} \\ &= 1 - \exp\left(\frac{r \log(1-c)I}{N} dt\right).\end{aligned}$$

Using a Taylor Series approximation then gives

$$\begin{aligned}dp &= 1 - \left(1 + \frac{r \log(1-c)I}{N} dt\right) \\ &= -\frac{r \log(1-c)I}{N} dt.\end{aligned}$$

Dividing by dt and taking the limit as dt goes to 0 gives the infinitesimal probability of a susceptible individual acquiring the disease:

$$\frac{dp}{dt} = -\frac{r \log(1-c)I}{N}.$$

In a short time dt , the number of susceptible individuals who become infected is the product of this quantity and the total number of susceptibles. Therefore, the equation governing the change in the size of the susceptible population is

$$\frac{dS}{dt} = \frac{r \log(1-c)SI}{N},$$

as desired.

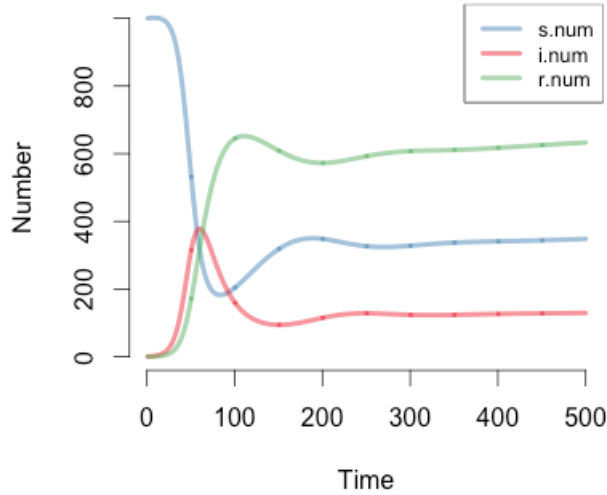
The best way to think of this equation is as representing flow out of the S compartment. This flow empties into the I compartment, which accounts for the first term in the second equation. The second term is due to the fact that individuals flow out of the I compartment as they recover, giving

$$\frac{dI}{dt} = -\frac{r \log(1-c)SI}{N} - \gamma I.$$

Of course, as infected individuals recover, they enter the R compartment, and we obtain

$$\frac{dR}{dt} = \gamma I.$$

FIGURE 1. Compartment sizes in SIR simulation.



We should point out that even in this very simple case, the differential equations are nonlinear, and must be solved numerically. Figure 1 shows plots of the relative sizes of the S , I , and R compartments in a typical scenario.

The Circular Migrations Model. The equations of the circular migrations model are similar to those of the SIR model in that they too are built around the concept of flows into and out of compartments, but here these compartments are greater in number and have a larger number of incoming and outgoing flows. There are six classes of people in the population: urban migrant males (MMU), rural migrant males (MMR), non-migrant urban males (NMU), non-migrant rural males (NMR), non-migrant urban females (NFU), and non-migrant rural females (NFR). Each of these is split further into four compartments: susceptible, acutely infected, chronically infected, and late-stage infected (denoted by S , A , C , and L , respectively). This gives a total of 24 compartments and equations in our model. The full set of equations can be found in [1], but we will examine one closely to get a feel for the mechanics of the model. Below is the equation governing the change in size of the susceptible urban migrant male population:

$$\begin{aligned} \frac{dS_{MMU}}{dt} = & \frac{\nu}{8} - S_{MMU}t_{MMU} \frac{A_{NFU}}{N_{NFU}} \beta_A - S_{MMU}t_{MMU} \frac{C_{NFU}}{N_{NFU}} \beta_C \\ & - S_{MMU}t_{MMU} \frac{L_{NFU}}{N_{NFU}} \beta_L - \delta S_{MMU} + \delta S_{MMR} - \mu S_{MMU}. \end{aligned}$$

The first term is for recruitment of new members into the population. The parameter ν is the overall “birth rate” across all compartments (it is convenient to use the term “birth”, but we should bear in mind that people enter the population not when they are born but when they reach sexual maturity). Each new member is equally likely to be a migrant male, non-migrant male, urban female, or rural female, and within the former two categories, equally likely to be urban or rural. Thus the recruitment rate for migrant males is $\frac{\nu}{8}$, which we see reflected in the equation.

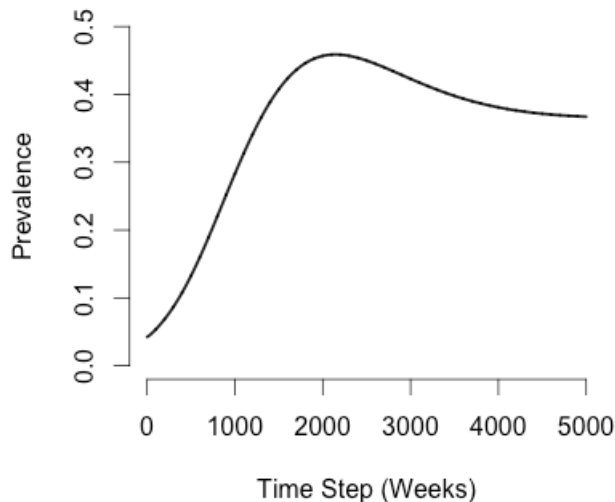
The next three terms describe the flow of individuals out of the S_{MMU} compartment and into the A_{MMU} compartment. The three terms correspond to the three different infection statuses of urban females. The chance that a particular susceptible migrant male is infected by an acute female is the product of the rate of sexual contact of migrant males, denoted t_{MMU} , the fraction of urban females who are acutely infected, $\frac{A_{NFU}}{N_{NFU}}$, and the per-act-transmission probability for acutely infected individuals, β_A . To obtain the expected number of susceptible MMU infected by acute females, we multiply these three factors by S_{MMU} . This is precisely the second term in the equation. The third and fourth terms arise in a similar way for infection of susceptible migrant males by chronic and late-stage females.

The fifth and sixth terms pertain to migration. The δ is the overall migration rate, so δS_{MMU} is the number of susceptible migrant urban males leaving for the rural location, and δS_{MMR} is the number of susceptible migrant rural males migrating into the urban location. The final term models mortality. The parameter μ is the overall death rate, so μS_{MMU} is the number of susceptible migrant males we expect to lose to death.

We should point out here that we have a choice regarding the parameterization of our model. We can either view sexual contact as occurring as isolated acts with potentially different members of the population (contact-as-act), or as occurring within extended partnerships (contact-as-partnership). In either case, the form of our equations remains the same, but some of the parameters take on different meanings. Instead of β_A representing the chance of transmission from an acute individual in the contact-as-act model, in the contact-as-partnership model it would represent the chance of transmission per time step from an acute partner to a susceptible partner. Similarly, instead of representing the frequency of sexual contact for migrant males, t_{MMU} would represent the frequency of partnership formation for members of this population. Details regarding the specific values of these parameters, which come from biological and behavioral data, and the relationship between the two characterizations of sexual contact can be found in [1].

Results. Models were programmed in the R programming language using the EpiModel package. The complete code can be found at <https://github.com/khanna7/circularmigrations/tree/master/epimodel/compartmental>. Figures 2 and 3 show the results from the contact-as-act and contact-as-partnership models.

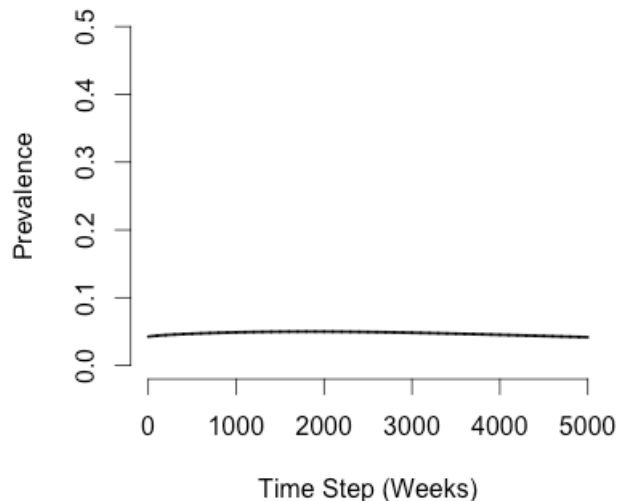
FIGURE 2. Prevalence from compartmental contact-as-act model.



It is easy to see why the prevalence in the former is quite high. Since none of the individuals in the population are part of a long-term sexual partnership, if a person becomes infected, he or she is immediately free to find a different partner to whom the disease could be transferred; there is no lag time during which a newly infected individual is less likely to infect someone else, as there is in the contact-as-partnership model while waiting for a "break-up." But why is the prevalence in the partnership model so low? One explanation is that because individuals aren't explicitly modeled, the influence of concurrent partnerships (a single person having more than one partnership at the same time) is suppressed. The result is that the only way in which infection status can change within a partnership is if one member infects the other. An example will help to illustrate the issue. Suppose that a susceptible man and a susceptible woman form a partnership. In the contact-as-partnership model, neither of them would get infected for the duration of the partnership, since both are healthy and can't give each other the disease. In reality, however, one of them, the man, say, could at the same time be in a second partnership with an infected member of the population. If he were to get the infection from this second partner, he could then transmit to the original partner. When we explicitly model members of the population in the network models below, concurrent partnerships can influence each other in precisely this way. But in the compartmental model, such interaction between multiple partnerships is not permitted, resulting in a lower observed prevalence.

The plots shown here are from simulations that used a 30 week migration interval. However, the analogous plots from simulations that used a 3 week migration interval were identical. In other words, migration interval has no effect on prevalence in the deterministic compartmental model paradigm.

FIGURE 3. Prevalence from compartmental contact-as-partnership model.



3. NETWORK MODELS

The deterministic compartmental models described in the previous section have the advantage of being both conceptually simple and easy to implement on a computer, but they lack several features that we ought to desire in a model of infectious disease transmission. One of these is a lack of stochasticity. Clearly, infectious disease transmission in a human population is a stochastic process, and we can be more confident in the results of a model that incorporates this random element at some level, especially because we can then use repeated simulation runs to generate confidence intervals and get a sense of the range of possible outcomes. A network model incorporates stochasticity by treating each node as a separate member of the population, with edges (representing sexual partnerships in our HIV model) between them forming randomly according to a pre-specified probability distribution. Another shortcoming of the differential equation models is that they do not capture the underlying dynamics of disease transmission. Infection, HIV or otherwise, spreads between individuals in a social network, and while this does create "flow" into and out of disease-status compartments, generating such flows by collecting data from a model of the lower-level transmission

network (where, as mentioned above, nodes are members of the population and edges represent contacts), is a closer approximation to reality. It is in fact possible to incorporate stochasticity into compartmental models, but such models still will not be detailed enough to model the underlying individual interactions that drive the infection.

Exponential Random Graphs. The most natural mathematical object to use in modeling a social network is the random graph, a collection of nodes among which edges are generated according to a given probability distribution. Because of the ease with which they can incorporate nodal attributes like sex or location and network configurations like triangles or stars (nodes with multiple edges) into the edge distribution, exponential random graphs (ERGs) have become the standard class of random graph used in the field of social network modeling, both epidemiological and otherwise. We explain the basic framework of these models below. We will have to make a few modifications for our specific application, but it is nevertheless essential to understand the basics.

Let S be a fixed set of nodes, and let Y be a graph-valued random variable taking values in the set G_S of all possible graphs (networks) on S . Let A be a collection of network configurations (for example, triangles, nodes with more than one edge, edges connecting nodes of two types in a bipartite graph, etc.), and for $y \in G_S$ and $a \in A$, let $g_a(y) = 1$ if the configuration a is present in y , and $g_a(y) = 0$ otherwise. The quantity $g_a(y)$ is called the network statistic of a , and henceforth we will write $a \in y$ to mean that a is present in y . We say that Y is an exponential random graph if the probability distribution of Y takes the form

$$P(Y = y) = \frac{1}{\kappa} \exp\left(\sum_{a \in A} \eta_a g_a(y)\right),$$

where $\{\eta_a\}_{a \in A}$ is a collection of constants (that can be either estimated or prescribed, depending on the particular application) and

$$\kappa = \sum_{y \in G_S} \exp\left(\sum_{a \in A} \eta_a g_a(y)\right)$$

is a normalization constant. The constants η_a can be thought of as weights representing the relative likelihood of observing each configuration $a \in A$. We can see this by examining the form of the probability distribution given above. For example, a large and positive η_a makes it more likely to observe the configuration a , since the exponential of a large positive number is itself large. On the other hand, if η_a has a large absolute value but is negative, then a is a relatively unlikely configuration. In the extreme case where $\eta_a = -\infty$, then for any $y \in G_S$ with $a \in y$, we have

$$\sum_{a' \in A} \eta_{a'} g_{a'}(y) = -\infty + \sum_{a' \in A \setminus \{a\}} \eta_{a'} g_{a'}(y) = -\infty,$$

so

$$P(Y = y) = \frac{1}{\kappa} \exp(-\infty) = 0,$$

that is, configuration a will never be observed. This is how we incorporate structural zeros into models; if we want to model a network of romantic relationships in a population consisting entirely of heterosexuals, we could let a be edges between nodes with the same sex, and then set $\eta_a = -\infty$ to eliminate the possibility of same-sex partnerships.

Could we possibly have $\eta_a = +\infty$, in order to insure that a certain configuration always occurs? In principle, we certainly could, but a closer examination reveals that the situation is a bit more complicated. If we allow some $\eta_a = +\infty$, then no finite value of κ could make the proposed formula a true probability distribution. In a specific application, one could conceivably work around this difficulty by assigning the required configuration separately, but we will not discuss this issue further since it is not relevant to the model at hand. We will simply say that for the purposes of this paper, the coefficients η_a are not allowed to take the value $+\infty$.

The drawback to this formulation of ERGs is that it is too “global” for our purposes. To simulate a network that evolves with time (as we wish to do), it is much easier to pass a computer a list of nodes and assign edges node by node than to sample from the distribution on G_S at each time step, which would be very expensive computationally, especially if A is large. Additionally, in a dynamic population of changing size, this distribution would be constantly changing, and would have to be re-computed at each time step.

We therefore need a “node-centric” formulation of ERGs. Suppose we are interested in the formation of edges from one time step to the next. More specifically, suppose i and j are two nodes, and that at time t , there is no edge connecting i and j . If we let $y_{ij,t} = 1$ if an edge exists between i and j at time t , and $y_{ij,t} = 0$ otherwise, then we are interested in the quantity

$$P(y_{ij,t+1} = 1 | y_{ij,t} = 0),$$

that is, the probability that an edge exists between i and j at time $t + 1$ given that no such edge existed at time t . We can compute this probability using the ERG formulation given above. An application of Bayes’ Formula (see [2]) yields,

$$\text{logit} (P(y_{ij,t+1} = 1 | y_{ij,t} = 0)) = \sum_{a \in A} \eta_a \delta^+(g_a(i, j)),$$

where $\text{logit}(p) = \log(\frac{p}{1-p})$, $0 < p < 1$, and $\delta^+(g_a(i, j))$ denotes the change in the value of $g_a(i, j)$ induced by adding an edge between vertices i and j to a graph in which there was none before. We can perform a similar calculation to obtain

$$\text{logit} (P(y_{ij,t+1} = 1 | y_{ij,t} = 1)) = \sum_{a \in A} \eta_a \delta^-(g_a(i, j))$$

for the probability of an edge persisting from one time step to the next.

From a computational point of view, these node-level edge formation and persistence laws are far more useful than the global formulation, and they are what we employ in our model; a new network is generated from the old network at each time step using these conditional probabilities. There is no reason that one must use the same model for both the formation and dissolution of ties. For example, we may be modeling a network for which we believe that although many factors contribute to the formation of partnerships, those partnerships, once formed, dissolve at a uniform rate. This is a common approach, and as we shall see it is the one we take here. In this case, the only configuration a we are concerned with in the dissolution model is the presence of an edge between nodes i and j , so that

$$\begin{aligned} \text{logit} (P(y_{ij,t+1} = 1 | y_{ij,t} = 1)) &= \eta_a \delta(g_a(i, j)) \\ &= \eta_a, \end{aligned}$$

since $\delta(g_a(i, j)) = 1$. In other words, the conditional probability of edge persistence is constant.

Circular Migration Network Model. In our particular model, we begin with a population of 5000 individuals divided into roughly the same classes as in the deterministic model (see previous sections for a refresher if needed). The only difference here is that we collapse MMU and MMR into a single MM class. Since each individual is explicitly modeled, there is no need from the point of view of the simulation engine to keep these two classes separate; people migrate according to the prescribed rate, and location tracking is taken care of automatically.

We design the simulation to be as similar to the contact-as-partnership model as possible. The formation of partnerships in the network is governed by an ERGM of the type described in the previous section. The main goal is to restrict partnership formation so that edges form only between heterosexual pairs and pairs in the same location. As mentioned above, we incorporate such “structural zeros” into the model by fixing the coefficients η_a of the relevant network statistics (such as the number of partnerships between MM and NFU) at $-\infty$. We should note here that stipulating that partnerships cannot form between two people in different locations is different from saying that such partnerships cannot exist at all; since a significant fraction of the population migrates at regular intervals, a partnership could form between a migrant man and a stationary female in the urban location, and then the male could depart for the rural location before dissolution of the partnership. In fact, this happens quite frequently, since the average duration of partnerships is 100 weeks, but migration occurs at 3 and 30 week intervals.

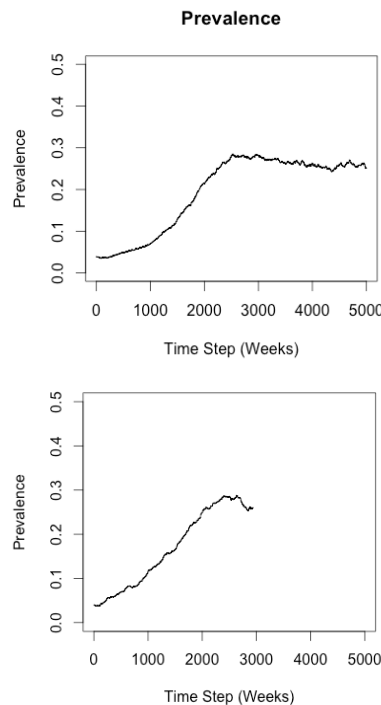
At each time step of the simulation, once partnerships are formed and dissolved (by applying the node-centric ERG formulations of the formation and dissolution models), infection is allowed to spread, individuals enter and exit the population (birth and death), and migrants change location. Infection is easy to model at this point; at each fixed time step, the network is nothing more than a fixed graph, and disease is transmitted along the edges according to pre-specified probabilities. These probabilities are the same as those used in the

contact as partnership model. One quirk to point out is that disease cannot be spread through a partnership consisting of two individuals in separate locations, for the obvious reason that the actors involved cannot have sexual contact at the current time step. In order to keep the average rate of sexual contact comparable across all members of the population, we stipulate that migrant males have on average twice as many partners as their non-migrant counterparts, since we expect a given migrant partnership to be "active" only half the time.

Birth rates, death rates, and migration rates are the same as those in the deterministic models, but in the network model the processes they govern are stochastic. For example, if ν is the overall birth rate (recall that "birth" really means "reaches sexual maturity" in this context), then where in the deterministic model $\nu/8$ is the fixed number of new MMU per time step, in the network model, the number of new MMU per time step is a Poisson random variable with mean $\nu/8$. And instead of a fixed fraction of migrants migrating at each time step, in the network model, each migrant is associated with a Bernoulli random variable with mean equal to the migration rate. If this variable takes on the value 1, then the migrant changes location. Otherwise, he stays put.

Results. As with the compartmental models, the network models were programmed in the R programming language using the EpiModel Package. See <https://github.com/khanna7/circularmigrations/tree/master/epimodel> for the complete code. Figure 4 shows the prevalence from simulations using 3 and 30 week migration intervals (technical difficulties halted the 3 week simulation at only 3000 steps, but this is enough to see the trend). The prevalence curves seen here lie in between those of the two compartmental models. This is as we expected, since the problems we identified with the contact-as-act and contact-as-partnership models that result in high and low prevalence, respectively, are fixed in the network models. Note also that each simulation had a similar peak prevalence occurring at around the same time step. To be able to state conclusively that migration rate has no effect on prevalence in the network models, we would have to run each simulation many times and generate confidence intervals for the prevalence at each migration frequency. Due to time and computing power limitations, this is not done in this paper (see [1] for a more thorough study), but the similarity between the prevalence curves in Figure 4 is evidence of the minimal effect.

FIGURE 4. Prevalence using migration frequency of 30 (top) and 3 (bottom) weeks.



Acknowledgments. It is a pleasure to thank my mentor, Aditya Khanna, for his guidance throughout the REU. Without the benefit of his expert advice this project would never have come together. I would also like to thank Nick Salter for his helpful discussions, and Peter May for organizing another great REU.

REFERENCES

- [1] A. Khanna, D. Dimitrov, and S. Goodreau, What can mathematical models tell us about the relationship between circular migrations and HIV transmission dynamics? *Mathematical Biosciences and Engineering*, **11** (2014), 1065-1090.
- [2] P. Krivitsky, M. Handcock, and M. Morris, Adjusting for network size and composition effects in exponential-family random graph models, *Stat Methodol.* **8** (2011), 319-339.
- [3] G. Robins, P. Pattison, Y. Kalish, D. Lusher, An introduction to exponential random graph (p^*) models for social networks, *Social Networks*, **29** (2007), 173-191.
- [4] M. Keeling and P. Rohani, Modeling Infectious Diseases in Humans and Animals, *Princeton University Press*, 2007.