Modeling of Social Contagion Project

Patrice Huang, Seoyoung Park

March 9, 2018

Abstract

This paper will show different forms of social contagion, obesity and emotions. Firstly, since an obesity epidemic has become one of the most concerning worldwide problems, in this paper, it is explained with an epidemiological model that considers both social contagion and non-contagious hazards of obesity, and compares several interventions' effectiveness as well as shows that the contagiousness of obesity decides the optimal choice of intervention programs. Secondly, to model emotional contagion, the SIS Model of Infection is used, but will also contain a different aspect that includes spontaneous contraction as well as the ability to contract the same emotions even after an infection. However, the study will show that despite the differences, emotional contagion still behaves similarly to infectious diseases.

1 Introduction

Contagion is the transmission of a disease by direct contact with an infected person or object, a disease or poison transmitted in this way. Social contagion is a subset of contagion like behavioral contagion, or a type of social influence. It refers to the idea of how something can spread and become popular, or the propensity a certain behavior is copied from one person to another. There are several types of social contagion, like hysterical contagion, financial contagion, and emotional contagion. In this paper, we will discuss social contagion and the effects of social contagion on two different models: obesity and emotions. We will also briefly discuss the earliest well-known model of contagion, the SIR model, as well as another model known as the SIS model.

2 Social Contagion of Obesity

2.1 Problem Description

As obesity has become one of the highest concerning global problems, a number of different interventions have been studied. Christakis and Fowler illustrated that obesity is able to spread out among individuals through a social contact network. The groundbreaking discovery of the spread of non-infectious disease via a social contact network was not only limited to obesity but also additional problems related to health like smoking.

Knowing that a part of obesity results from person-to-person transmission, the effectiveness of significant control programs against the spread of obesity can be expressed by nonlinear dynamics with a connected risk. The danger of obesity is identified by not only a particular person but also other people in the same population. The dependence suggests that herd effect by an implement of public health interventions can be assumed, but it also suggests that the contagious effect could result in social problems such as social discrimination and friendship network.

To compare the effectiveness of several interventions, an epidemiological model is constructed to indicate the time-dependent and age-dependent risk of obesity, with the threat treated as both dependent and independent of obesity prevalence, using observed data.

2.2 Mathematical Model

2.2.1 SIR Model of Infection

This model was created by Lowell Reed and Wade Hampton Frost in the 1920s. The dependent variables are put into three classes: susceptible to some contagious phenomenon(S), infective(I), and Recovered(R). Each are a function of time:

S = S(t) is the number of susceptible individuals(never-obese)

I = I(t) is the number of infected individuals(obese)

R = R(t) is the number of recovered individuals(ex-obese)

The second set of equations represents the fraction of the total population in each of the three classes, with N as the total population.

$s(t) = \frac{S(t)}{N}$ is the susceptible fraction of the population

$i(t) = \frac{I(t)}{N}$ is the infected fraction of the population

$r(t) = \frac{R(t)}{N}$ is the recovered fraction of the population

β and γ are positive constants. The total population (S+I+R) is constant. The three differential equations that represent the SIR model are:

$$\frac{dS}{dt} = -\beta IS \tag{1}$$

$$\frac{dI}{dt} = \beta IS - \gamma I \tag{2}$$

$$\frac{dR}{dt} = \gamma I \tag{3}$$

2.2.2 A Model for the Social Contagion of Obesity

We can express the epidemiological process of becoming and recovering from obesity as a function of time, supposing obesity is resulted from both contagious and non-contagious routes.

$$\frac{dS}{dt} = \mu N - [\beta I(t) + \varepsilon]S(t) - \mu S(t) \tag{4}$$

$$\frac{dI}{dt} = [\beta I(t) + \varepsilon]S(t) + \sigma[\beta I(t) + \varepsilon]R(t) - (\mu + \gamma)I(t) \tag{5}$$

$$\frac{dR}{dt} = \gamma I(t) - \sigma[\beta I(t) + \varepsilon]R(t) - \mu R(t) \tag{6}$$

N: the total population size (constant over time), $N = S(t) + I(t) + R(t)$ for any t

μ: the birth and death rate of human host

β: the transmission coefficient

ε: the hazard of obesity by non-contagious reasons

γ: the natural recovery rate

σ: the relative risk of weight regain among ex-obese individuals that especially takes greater than

1 because of high risk of becoming the obese state back

In this system, we consider that ex-obese(R) is not contagious. From never-obese people(S), the hazard rate of obesity when they face the obesity for the first time, also known as "force of infection", is $\lambda(t) = \beta I(t) + \varepsilon$. As shown in the equation, the hazard of obesity for ex-obese(R), is σ times greater than that for never-obese(S).

In the hazard rate of obesity, two hazards are added: one is through the contagious route $\lambda_1 = \beta I(t)$, and the other is through the non-contagious route $\lambda_2 = \varepsilon$ which is constant. This implies that the contagious and non-contagious hazards are independent; however, it would be more reasonable to consider dependence between β and ε, considering that the social contagion has to gradually affect dietary behavior and physical activity to reach a "transmission of obesity" in human life. The result depends on the transmission coefficient of obesity and a restricted transmission that helps prevent weight gain among never-obese people.

2.3 Solution of the Mathematical Problem

It is obvious that the initial condition for $S(0)$ is N, the total population. Solving equations (4), (5), (6), equilibria S^*, I^*, R^* are found to be asymptotically stable so that all the system's trajectories converge.

We are going to focus just on the age-dependent dynamics, not time-evolution, since a lot of epidemiological measurement of obesity for an individual is the hazard of obesity or related caused by diseases by a certain age, and we tend to understand the model for the social contagion of obesity as it is directly affected by the age-dependent risk of obesity. Here, we set the equilibria S^*, I^*, R^* in the age-specific risk with $X(a), Y(a), Z(a)$ where a is the age.

$$\frac{dX}{da} = -(\lambda^* + \mu)X(a) \tag{7}$$

$$\frac{dY}{da} = \lambda^* X(a) + \sigma\lambda^* Z(a) - (\mu + \gamma)Y(a) \tag{8}$$

$$\frac{dZ}{da} = -(\sigma\lambda^* + \mu)Z(a) + \gamma Y(a) \tag{9}$$

Here, the total population size N_c at age a is $X(a) + Y(a) + Z(a)$, and $N_c(a)$ is parameterized as

$$N_c(a) = N_c(0)e^{-\mu a} \tag{10}$$

Also, because the total population size is fixed over time,

$$N = \int_0^\infty N_c(0)e^{-\mu a}da = \frac{N_c(0)}{\mu} \tag{11}$$

$N_c(0)$ is equal to μN. Since people are assumed to be born as never-obese, the initial condition $(X, Y, Z) = (N_c(0), 0, 0)$, and $X(a)$ can be written as

$$X(a) = N_c(0)e^{-(\lambda^* + \mu)a} \tag{12}$$

We set the life-time risk as a probability of not staying in the never-obese situation through an individual's whole life, using $x(a) = \frac{X(a)}{N_c(0)}$, the probability to remain in the never-obese state by age a. The overall risk by age a, $q(a)$, is

$$q(a) = \int_0^a \lambda^* x(s)ds = \frac{\lambda^*}{\lambda^* + \mu}[1 - e^{-(\lambda^* + \mu)a}] \tag{13}$$

As a goes to ∞, $q(a)$ becomes $\frac{\lambda^*}{\lambda^* + \mu}$, which tells that the larger the prevalence, the larger risk for people to have obesity in their life.

Referring to published empirical data, we set values to parameters for equation (4),(5),(6):

Description	**Notation**	**Baseline value**
Population size	N	100,000
Average life expectancy at birth	$1/\mu$	69.4 (years)
Transmission coefficient of obesity	β	2.96×10^{-7} (per year)
Non-contagious hazard of obesity	ε	0.012 (per year)
Relative hazard of obesity among ex-obese	σ	8.0
Average duration of obesity	$1/\gamma$	35.8 (years)

The empirical household secondary attack proportion, SAP, with a single index case is constructed with the basic reproduction number, R_0 by

$$SAP = \frac{R_0}{R_0 + m} \tag{14}$$

where m is the number of susceptible and exposed people in the household. R_0 is derived mathematically from linearized systems in similar epidemic models. The disease-free equilibrium is always unstable for the system of the social contagion of obesity except when there is no non-contagious hazard of obesity(when $\varepsilon = 0$). Therefore, to make the non-contagious hazard of obesity equal to zero, we use $R_0 = \frac{\beta N}{(\gamma+\mu)}$. Let $m = 3$ and SAP ranged from 0.135 to 0.254, then the transmission coefficient β is 2.96×10^{-7} (the mid-point of 1.99×10^{-7} and 4.33×10^{-7}).

2.4 Results and Discussion

Using the baseline setting in table above, as time goes by, the prevalence converges to a stationary value. As shown in Figure 1A, the equilibrium state is reached after roughly 200 years, and the prevalence is calculated at 60.8%. Figure 1B shows the age distribution of S,I, and R in terms of age a by using the age-dependence system, and the highest risk of obesity is calculated to occur at the age of 37 years, and after that the risk decreases because of natural mortality.

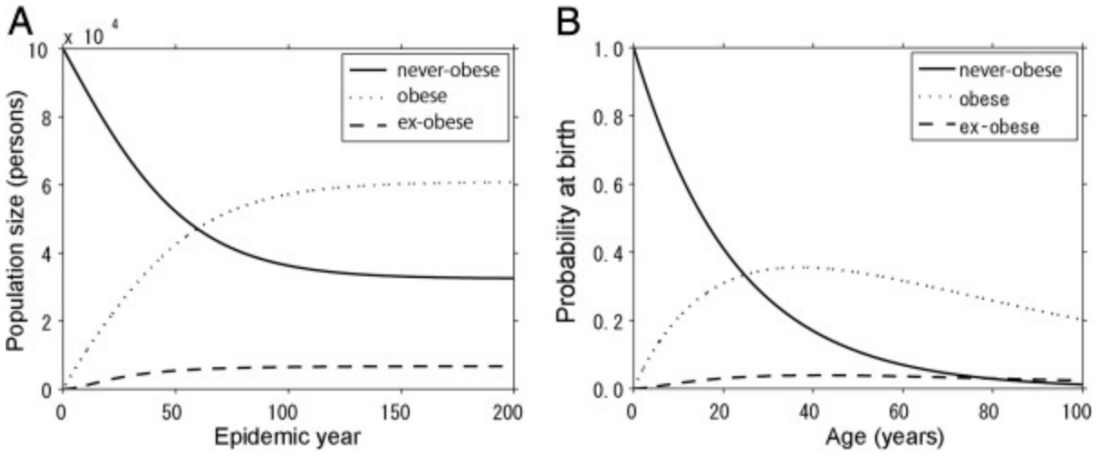

Figure 1

The two primary prevention methods are the intervention on social contact, which affects

the transmission coefficient β, and the preventing weight gain among the never-obese population, which affects the non-contagious hazard ε. The intervention of social contact's primary function is to repress the contagion by averting social transmission between people, while the second prevention method is to control the diet and restrict calorie consumption, as well as increase physical activity. As shown in Figure2A, as β gets higher, the prevalence get higher as well, but the equilibrium prevalence seems sensitive to β and varying at some value of β depending on ε. It also shows that the prevalence goes backward, which means an endemic equilibrium exists for $R_0 < 1$, demonstrating that it would be hard to control obesity in the presence of person-to-person transmission. The two types of interventions in the secondary prevention, which are controlling the diet among obese people and following up on ex-obese people, each have different effects, which are the duration of obesity $1/\gamma$ and the relative hazard among ex-obese σ respectively. From FigureB, as $1/\gamma$ gets shorter, the equilibrium prevalence becomes shorter, and varying σ makes the prevalence less sensitive to $1/\gamma$.

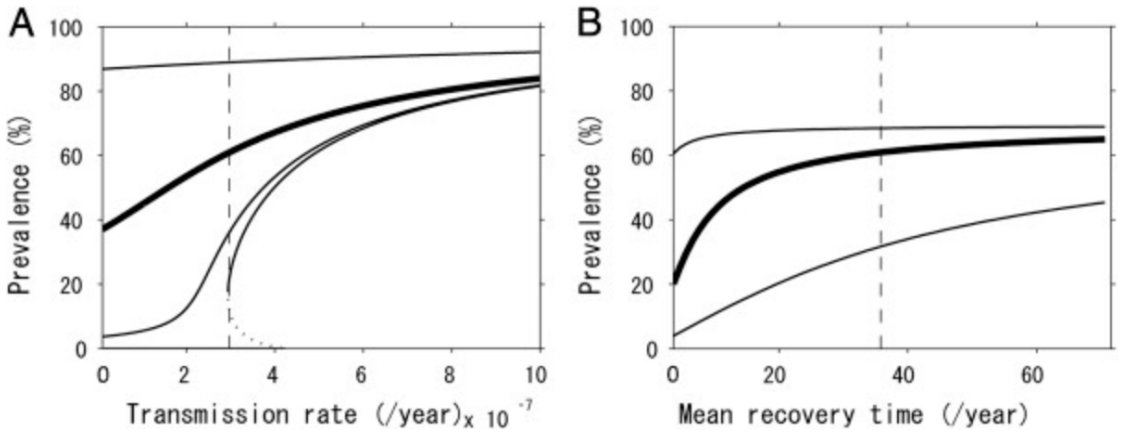

Figure 2: A. The bold line shows the baseline result depends only on the transmission coefficient β. The others show the cases when ε is varied to the 0%, 10% and 1000% relative to the baseline value (from the horizontal axis to the top). B. The bold line shows the equilibrium prevalence of obesity using baseline parameter values other than the average duration of obesity, $1/\gamma$. Two other lines show the cases when σ is set to 10% (bottom) and 1000% (top), respectively, relative to the baseline value.

To compare several control programs, interventions in the primary and secondary preventions are examined separately. The parameter set that controls the primary prevention is α, and it is assumed both contagious and non-contagious risks get lower by α.

$$\frac{dS}{dt} = \mu N - \alpha[\beta I(t) + \varepsilon]S(t) - \mu S(t) \tag{15}$$

$$\frac{dI}{dt} = \alpha[\beta I(t) + \varepsilon]S(t) + \sigma[\beta I(t) + \varepsilon]R(t) - (\mu + \gamma)I(t) \tag{16}$$

$$\frac{dR}{dt} = \gamma I(t) - \sigma[\beta I(t) + \varepsilon]R(t) - \mu R(t) \tag{17}$$

As you notice, only never-obese people's hazards get reduced. Since reducing prevalence by α in the presence of dependence can be high, a predicted independence between β and ε might make it overestimate the effectiveness of primary prevention.

For secondary prevention, the factor enhancing programs is set as κ.

$$\frac{dS}{dt} = \mu N - [\beta I(t) + \varepsilon]S(t) - \mu S(t) \tag{18}$$

$$\frac{dI}{dt} = [\beta I(t) + \varepsilon]S(t) + \kappa\sigma[\beta I(t) + \varepsilon]R(t) - (\mu + \kappa\gamma)I(t) \tag{19}$$

$$\frac{dR}{dt} = \kappa\gamma I(t) - \kappa\sigma[\beta I(t) + \varepsilon]R(t) - \mu R(t) \tag{20}$$

Notice the overall hazard of re-infection is reduced in the program.

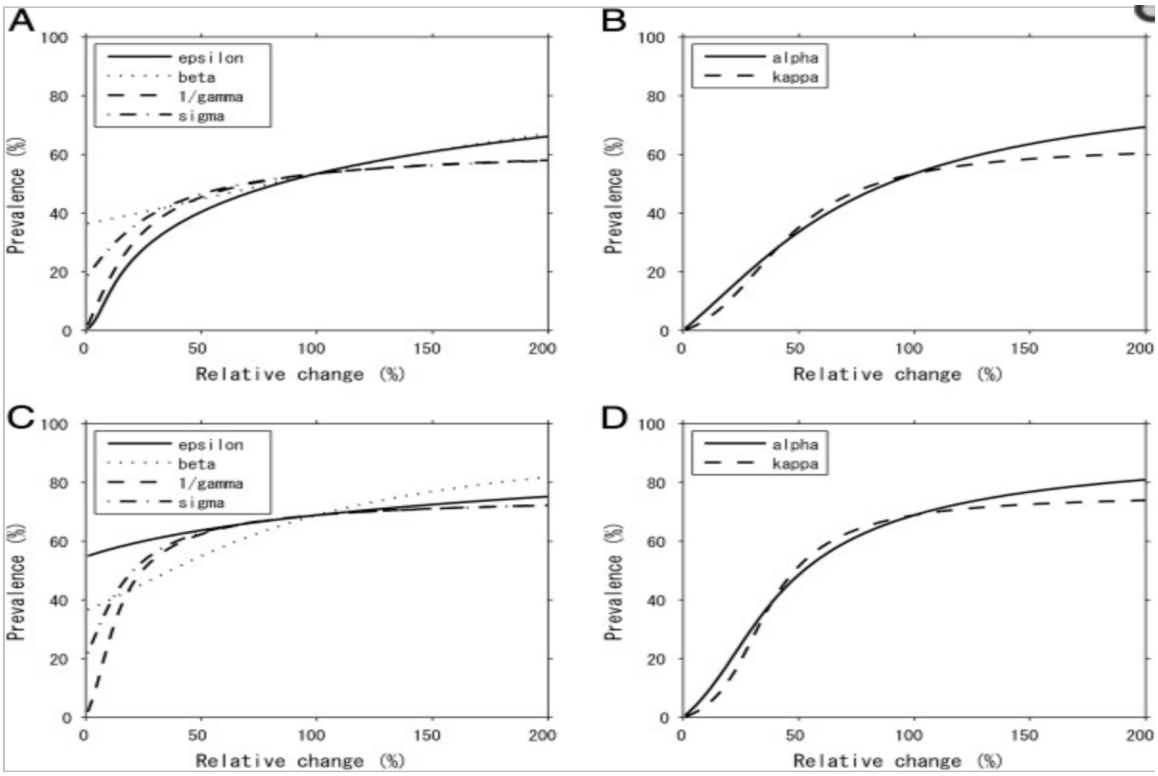

Figure 3: A&C shows the prevalence when single parameters (ε,β, $1/\gamma$ and σ) are independently varied. B&D compare the primary prevention (reducing α) and the secondary prevention (reducing κ). For A &B, β is set to be low($1.99 \times$ 10-7 per year), while β is set to be high for C & D($4.33 \times$ 10-7 per year).

In the Figure 3A, when β is small, ε is the lowest, which means preventing weight gain among never-obese people is most effective. $1/\gamma$, dietary restricting among obese people, is the second effective. In other words, while β stays at a very small value, intervention that plans for reduction of the non-contagious hazard would be the most effective, and rapidly getting rid of obese people by the program would be assumed to reduce obesity effectively. In the Figure 3B, targeting never-obese individuals or targeting obese and ex-obese individuals to reduce overall hazards of obesity is similarly effective.

When the transmission coefficient is high to maintain the transmission of obesity, in Figure 3C, ε(preventing weight gain among never-obese individuals) seems to be the least effective, while $1/\gamma$(the dietary regulation) and σ(implementing the follow-up program) are the most effective. If we can select two strategies to combine, the primary and secondary preventions result in similar impacts, and the effectiveness relies on the power of the interventions at an individual level.

2.5 Improvement

There are five things to consider as restrictions for this study:

1. The model uses a homogeneously mixing assumption for simplicity, so to improve the study, it should take into account more appropriate social contact network for later studies.

2. Because of the lack of history of obesity, age-dependent heterogeneity is not able to be considered here. It would be better with more information.

3. The equilibrium prevalence is calculated higher(60.8%) than current prevalence in the real world.

4. For efficiency, the model doesn't compare the cost of intervention programs.

5. Further work could regard other choice of modeling. For example, the conditional risk model with stochastic dependence structure between risks with different routes of obesity could have helped the problem of overestimation of the primary prevention's efficiency.

2.6 Conclusions

Using the system of the social contagion of obesity, choosing interventions to reduce obesity depends on the transmission potential of obesity between individuals. It is crucial to quantify the epidemiological dynamics of obesity in advance to get proper assessment and comparison of variety of public health control programs of obesity.

3 Emotions and Social Contagion

Social network is a network of social interactions and personal relationships. Social networks have great effects on human behavior. Since people interact with lots of different people, their social interactions and how often they interact with certain people affect their behaviors. Research in lots of different fields have shown that social networks affect people in almost every way, from economic behavior like purchasing decisions and bankruptcy, to college GPA, crime, and even things like smoking, suicidal ideation, and cooperative behavior evolution.

3.1 Problem Description

"Emotional contagion" is a type of social contagion, and is the influence of human emotions by social contacts. Through previous studies, it has been shown that positive and negative emotions can occur over small timescales and through close contact with others, or through media. This study shows the interpersonal spread of long-lasting emotional states over a large period of time, rather than short periods of time.

Through other studies by Fowler & Christakis, there has been evidence of happiness in relation to social contracts, which is extracted from the Framingham Heart Study. The study showed that there is correlation between happiness and social networks, which says that if your friends are happy and you are often with your friends, you have a greater chance of being happy in the future. This shows an inductive effect of happiness spreading from person to person, which shows happiness as a form of social infection. Mathematical models for the spread of microbial infections

have helped others model the spread of disease, as well as other non-microbial infections like rumors or computer viruses.

3.2 Simplifications

This particular study is showing a new approach for studying the spread of emotions in a social network. Emotions have three levels: moods, states, and traits. Each level shows an increase in duration and permanence, and therefore have different levels of mutability of immutability. The terms "content" and "discontent" will be synonyms for happiness and depression, with emotions based on long-term emotional states. This study will focus on the SISa Model of Infection, which is a modification of a classical mathematical model to account for differences between disease and emotion. The Framingham Heart Study dataset of emotional states will be what this study uses, to see if emotional states act like infectious diseases.

The criteria an emotional state (like content) need to be similar to an infectious disease is:

1. the probability of becoming content (or any specific emotional state) must depend on the number of contacts (people in the social network) that are content

2. the probability of switching back from content to neutral ("recovery") should be independent of any properties of social contact

3.3 Mathematical Model

3.3.1 SIS Model of Infection

Individuals are either susceptible, which means that they do not have the disease, or infected, which means that they do have the disease. A susceptible person can get the disease if they come in contact with an infected person. The rate of disease transmission from an infected person to susceptible person is given by the variable β, which is also known as the transmission rate. After someone is infected, they recover from the disease at the constant rate γ, regardless of their contact with other susceptible or infected individuals.

The SIR Model of Infection is one class of disease models, and is also one of the simplest infectious

disease models. The SIR model is the susceptible-infected-recovered model, where individuals that are infected eventually recover, and become immune to further infection and enter a "recovered" state.

For emotions, the SIS Model of Infection is used, which is the susceptible-infected-susceptible model. This is because an emotion can reoccur in someone's life many times, rather than the infection happening once and then the individual becoming immune to the disease. Therefore, after the infection, it is assumed that the individual goes back to being susceptible rather than "recovered". The SIS Model is often used for infectious diseases that do not have immunity, including several sexually transmitted diseases.

3.3.2 SISa Model of Infection

The standard SIS Model is different from the SISa Model of Infection. The SIS Model only counts infection if a susceptible individual comes in contact with an infected individual. However, emotions are different from infection, and can occur through other ways rather than just physical contact. This is where the "a" in the "SISa" Model comes from, which is an extended factor for where uninfected individuals automatically become infected at the constant rate of α, which is independent of infected contacts.

The SISa Model is defined with a set of differential equations, which helps model a well-mixed population.

$$\frac{\partial S}{\partial t} = -\beta IS + \gamma I - \alpha S \tag{21}$$

$$\frac{\partial I}{\partial t} = \beta IS - \gamma I + \alpha S \tag{22}$$

The variables are each a function of time, where:

S = S(t) is the number of susceptible individuals

I = I(t) is the number of infected individuals

and

$$I + S = N \tag{23}$$

β is the transmission rate, γ is the recovery rate, and α is the rate of spontaneous infection. This model assumes that the population stays constant.

In the SISa model, there are 3 possible processes where an individual's state can change, which are shown in Figure 4A.

1. An infected individual transmits the disease to a susceptible individual at rate β
2. A susceptible individual spontaneously becomes infected at rate α
3. An infected individual goes back to being susceptible at rate γ

The rate of movement (Figure 4B) for an individual between the susceptible and infected states are:

1. The rate from susceptible to infected is $\alpha + \beta n_1$
2. The rate from infected to susceptible is γ

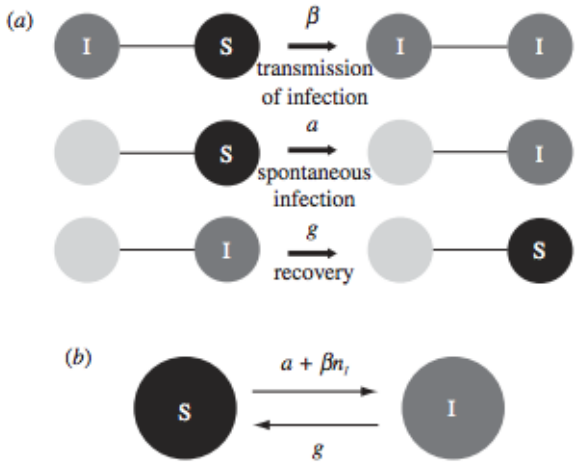

Figure 4: Figure A shows the three processes by which as individual's state can change, while Figure B shows that rates of movement for an individual between the susceptible and infected states.

3.4 Solution of the Problem

The classical definition of an infectious disease in the SIS Model is:

1. the probability of an individual transitioning from susceptible to infected is an increasing function of the number of infected contacts

2. the probability of a transition from infected to susceptible (recovery) is independent of the number or state of contacts

The two ways to become sick is if you are exposed to infected individuals, or through spontaneous disease acquiring. Using information about transitions between the different states on multiple time points in the data separated by Δt, the probability of a transition from susceptible to infected after time Δt is about $(\alpha + \beta n_1)\Delta$t, and the probability of the transition from infected to susceptible is $\gamma\Delta$t.

The data used is from the Framingham Heart Study. Background on this data is necessary to understand how the data is collected. The study is continuous, with subjects being examined at regular, 3 year intervals. The subjects are given a test called CES-D, where each individual is classified as content, discontent, or neutral. Social networks were also constructed for each subject based off of information given on family, friends, coworkers, and neighbors. In this study, the neutral state corresponds to the susceptible state, and content and discontent are infected states. Although coinfection is not possible, superinfection is possible (spontaneous infection, from content directly to discontent or vice versa). The data is then used to calculate values for the parameters as well as validate the model's assumptions.

A state x is infectious if having more contacts that are in state x makes an individual more likely to switch to state x. For example, a positive relationship between the number of contacts in state x and the probability of transitioning from state y to state x shows that in respect to state y, state x is infectious.

To determine whether a state x is infectious with respect to state y, an ordinary least squares linear regression is performed. There are two exams, exam 6 and exam 7. Each subject in state y that is in exam 6 is given a code, as either transitioned to state x (transition = 1) or not (transition

= 0). Then, using regression for each subject against the subject's number of contacts in state x during exam 6, there is either a positive or negative correlation. A significant positive correlation means that having more contacts in state x will make you more likely to switch to state x later. If there is a positive correlation, it means state x is infectious and β can be calculated from the slope of the regression line, while α can be calculated from the intercept. If there is no significant correlation, it shows that state x is not infectious, and γ can be calculated from the intercept. The data is only valid if the subjects do not change emotional states rapidly, but rather changes by the year or even longer.

This approach to social contagion is different from other models of correlation in the past. This approach looks at how contacts influence the transition between different states, which captures contagion more accurately. There is no selection bias, and confounding events do not show up as contagion effects in the model.

3.5 Results

Using the data of individuals who have been administered the CES-D, there is a population size of n=1880 and an average degree of 3.1. However, this network had Poisson distribution, with a long tail, and showed that subjects with friends or family living 2 or more miles away did not have a significant association with the subject. Therefore, the analysis restricted connections to within 2 miles, while changed the degree of the network to 1.25 ± 1.35.

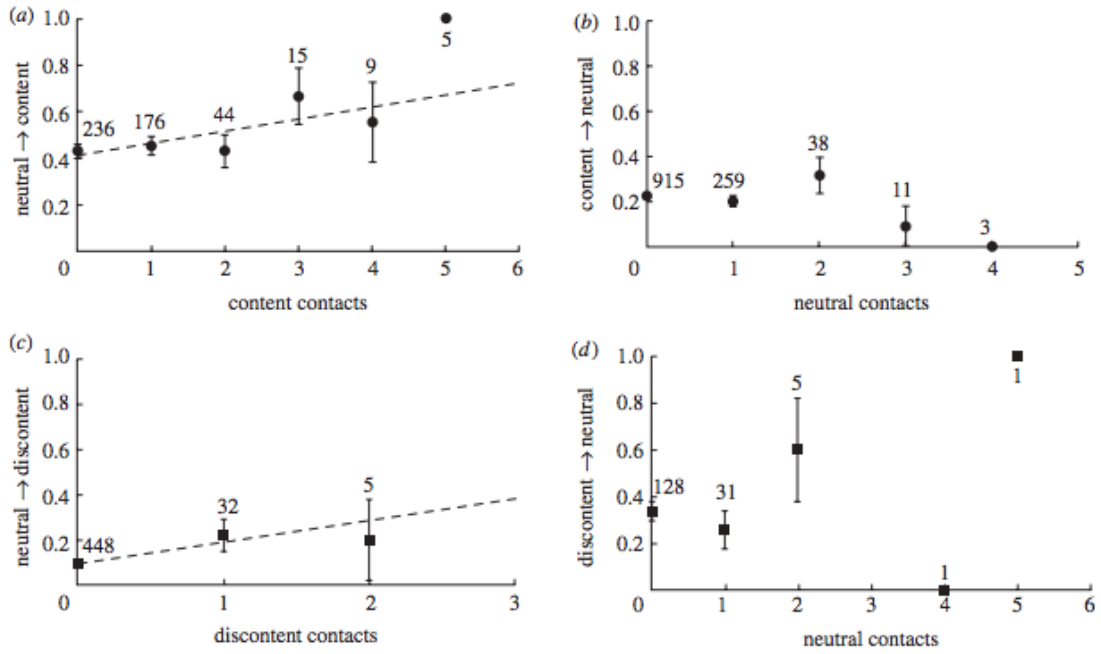

Figure 5: Content and discontent both behave like disease agents, and infect those in a susceptible neutral emotion state. Figure A shows that the probability of transitioning from neutral to content increases in the number of content contacts. Figure B shows the probability of recovering from content to neutral does not depend on number of neutral contacts. Figure C shows that the probability of transitioning from neutral to discontent increases in the number of discontent contacts, and Figure D shows that the probability of recovering from discontent to neutral does not depend on neutral contacts.

Figure 5A through D show the results of the infectiousness analysis for the spread of content and discontent. The results are consistent with the SISa model, as they show a significant positive correlation between the probability of transitioning from neutral to content and the number of content contacts (Figure 5A), and no significant relationship between the transition from content to neutral and the number of neutral contacts (Figure 5B). As there was no significant relationship between the probability of transitioning from neutral to content and the number of neutral contacts, or transitioning from content to neutral and the number of content contacts, it shows that content can be modelled as an infectious process in the SISa framework.

Discontent can also be described with the SISa framework as a contagious emotion with neutral as the susceptible state, with the probability of transitioning from neutral to discontent being positively correlated with the number of discontent contacts (Figure 5C) and the probability transitioning from discontent back to neutral being independent from the number of neutral contacts (Figure 5D). There was also no significant relationship between the probability of transitioning

from neutral to discontent and the number of neutral contacts, or transitioning from discontent to neutral and the number of discontent contacts.

The rate of transmission through contact is similar for content and discontent, but the rate of spontaneous infection is much higher for content and the rate of recovery is higher for discontent.

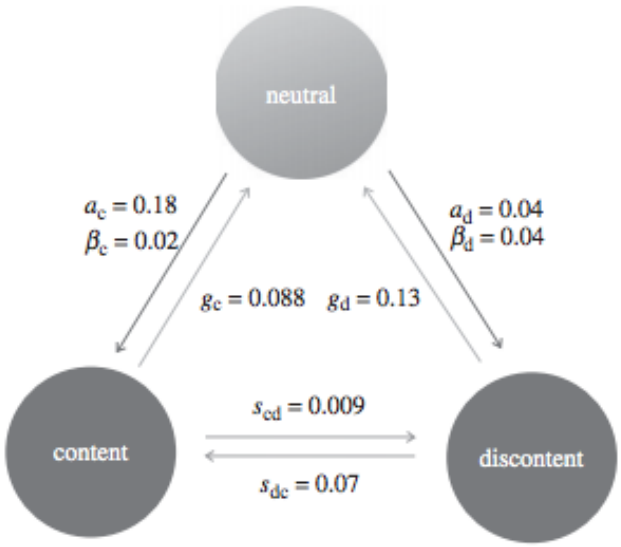

Figure 6: This figure shows the parameter estimates for content and discontent using the SISa model framework, which were derived from Figure 5. These results show that content and discontent are infectious emotions that can be 'caught' by susceptible (neutral) individuals from infected (content or discontent) contacts. Superinfection is also shown with the transfer directly from content to discontent or vice versa. α_c and α_d are the rate of spontaneously becoming content or discontent, γ_c and γ_d (g_c and g_d) are the rate of recovery from content or discontent, β_c and β_d are the transmission rates of content and discontent, and s_c and s_d are the rate of spontaneous superinfection from content to discontent or discontent to content.

Figure 6 shows all the calculated parameters from the SISa model and the data used. The average lifetime of an emotional state is the average length of time the individual spends in that state before recovering (often going to the neutral state (γ_c or γ_d, the rate of recovery from content or discontent), and also includes the possibility of superinfection (going directly from discontent to content or vice versa) (s_{cd} or s_{dc}). The sum of the two rates is the total recovery rate (γ_T). The average lifetime of contentment infection is 10.1 years, but for discontentment it is 5.0 years, half of the contentment lifespan.

The "influence" of an emotion is the cumulative probability that the infection will be passed from an infected to susceptible before the individual recovers from the infection. This is given by 1 - $e^{-\beta/\gamma_T} \simeq \beta/\gamma_T$. This value is 0.18 for both discontent and content. The 'cycle length' is the

average length of time between spontaneous infections, and is 5.6 year for content and 25 years for discontent, meaning an individual spontaneously becomes content more times than discontent.

3.6 Discussion

In a regular SIS model, there is the basic reproductive ratio, R_0, which is the average number of secondary infections caused before recovery. An epidemic could only occur for when $R_0 > 1$. However, in the SISa model, there is no thresholding behavior because of the spontaneous infection term, as even with no initial infections or contagion, the spontaneous term will cause infection. Without the spontaneous infection, R_0 would be estimated as $\beta \mathrm{n}/\gamma$, with n as the average degree. For both content and discontent, R_0 would be less than 1, meaning neither would have self-sustained epidemics, and neither emotion would be sustained without spontaneous infection.

Further analysis also showed that individuals in different states have different network structures than discontent individuals, with content individuals having on average 1.26 connections to others, neutral people having 1.37 connections and discontent individuals having 0.97 connections. This shows that discontent individuals tend to stay more in isolation, which could be a contribution to their depressive symptoms. Clustering also change with people in different states.

	content	discontent	neutral
average degree	1.26	0.97	1.37
observed clustering	$C_{\mathrm{cc}} = 1.04$	$C_{\mathrm{dd}} = 1.07$	—
	$C_{\mathrm{cn}} = 0.91$	$C_{\mathrm{dn}} = 0.77$	—
simulated clustering	$C_{\mathrm{cc}} = 1.09$	$C_{\mathrm{dd}} = 1.92$	—
	$C_{\mathrm{cn}} = 0.95$	$C_{\mathrm{dn}} = 1.0$	—

Figure 7: This table shows the network structure dependence on emotional states, and comparison with epidemic simulations. The discontent individuals have a lower number of contacts, shown by the lower average degree, than neutral and content. C_xy is the clustering between individuals of type x and y, with c as content, d as discontent, and n as neutral.

Clustering is the ratio of the observed number of connections between two types of individuals to the number of connections expected if the positioning of individuals was random. In a infection network, the average correlation between two infected individuals was greater than one, while the average correlation between an infected and susceptible individual is less than one. This can be

explained by inductive spread of content between content individuals. For discontent individuals, there was less clustering in the data than predicted by the infected individual simulations, which supports that discontent people have less clustering, which could be explained by the tendency for depression to cause people to interact less with others.

Through the study, it is shown that long-term emotional states can spread between socially connected individuals. Due to the transmission rates (β and α) found through regression, the probability per year of becoming content through each content contact increases by 11%, and each discontent contact increases the individuals chance of becoming discontent by 100% each year.

This framework created for formalizing social contagion derived from study of infectious diseases can be used in the study of the spread of other emotions or other social network phenomena. This approach has advantages over other methods, as this method controls selection bias as well as confounding environmental factors that affect contacts. The probability of changing emotional state also does not have any dependence on age, gender, or education, which have been factors of dependence in other studies in the past. The lack of dependence on such factors shows that this approach is better in including variation across individuals than other methods.

3.7 Improvement

A possible limitation in the study is the dataset used for social network. Using the Framingham Heart Study means that the data in that study is assumed to be the same for the spread of emotions overall, which is untrue. Important contacts could've also been missed in the study. This doesn't particularly change the results qualitatively, but it changes the parameters and rate constants. Different network connections could also have different weight on their ability to transmit emotion. Ideally, a longitudinal study would be done measuring social networks and health, as well as defining specific contacts by the amount of time spent together and their influence.

Another possible point of improvement is the definition of "discontent". A CES-D score above 16 can occur if someone has major depression. Therefore, in the study, people with depression were included, but it also assumed that the probability of becoming discontent is independent of the

individual's history of depression. Therefore, this also could have an effect on the data.

The study shows the transmission component of the emotions of content and discontent. It has been shown that short-term emotional states can be coordinated quickly and through direct contact, but it is not obvious how the contagion behaves for long-term emotional states. The observation that 'catching' content and discontent happen through social contacts but 'recovering' to a neutral state doesn't shows that infection is similar to teaching a life skill that lasts for a long time without having constant stimulation. To find the truth behind long-term emotional states, more study is required.

3.8 Conclusions

This study demonstrated that emotions can be modelled, and introduces an upgraded model that has several advantages over other models. This model can study content and discontent, and gives information on the transmissive nature of positive and negative emotions as well as other emotions.

Using the SISa Model of Infection to account for emotions contracted spontaneously and through transmission, the emotional states in the Framingham Heart Study dataset show that over long timescale, both content and discontent behaviors are similar to the behaviors of infectious diseases.

References

Ejima, K., Aihara, K., & Nishiura, H. (2013). Modeling the obesity epidemic: social contagion and its implications for control. Theoretical Biology & Medical Modelling, 10, 17. http://doi.org/10.1186/1742-4682-10-17

Hill, A.L., Rand, D. G., Nowak, M.A., & Christakis, N. A. (2010) Emotions as infectious diseases in a large social network: the SISa model. *Proceedings of the Royal Society B, 277,* 3827–3835. http://rspb.royalsocietypublishing.org/

Martcheva, M. (2015). An Introduction to Mathematical Epidemiology. *Springer*, New York. Vol 61. https://link.springer.com/content/pdf/10.1007/978-1-4899-7612-3.pdf