Modeling Disease Spread and Health Behavior Diffusion through Clustering and Social Network Exposure

Divya Aggarwal[*]
Humanities and Social Sciences Department
Indian Institute of Technology Roorkee
Roorkee, Uttarakhand, India
d_aggarwal@hs.iitr.ac.in

Abstract

This report investigates the dynamics of infectious disease transmission, emphasizing the roles of social interactions, demographic factors, and social interactions. By incorporating the concept of Social Network Exposure (SNE) into traditional epidemiological models, the report reveals how weak ties among diverse socioeconomic groups can influence disease spread. It highlights the significance of age-structured approaches in understanding susceptibility and recovery patterns across different cohorts. Additionally, the integration of latent and asymptomatic infection stages captures the effects of social ties on disease dynamics within closely connected communities. The findings underscore the impact of clustered social networks on prolonging infections and sustaining transmission. Collectively, these insights provide a comprehensive understanding of the multifactorial nature of disease spread and inform public health strategies to mitigate health disparities and improve health outcomes across populations.
Keywords: Infectious Disease Dynamics; Social Network Exposure; Clustering Coefficient

1 Introduction

The spread of infectious diseases and the persistence of chronic health conditions are complex phenomena shaped by numerous factors, such as socioeconomic status, social connections, and individual health behaviors [1], [2], and [3]. Understanding these influences is essential for crafting impactful public health interventions, especially given current global health challenges. This report presents a comprehensive exploration of these dynamics through four distinct analyses, each addressing a critical aspect of disease transmission.

Infectious diseases and chronic health conditions spread through both biological and social pathways. Traditional epidemiological models, like the Susceptible-Infectious-Recovered (SIR) framework [4], have long been used to predict disease spread. However, these models often fail to account for the complexity of social networks that influence not only the rate of transmission but also the reach of health behaviors across socioeconomic classes.

The motivation behind this study is to extend these classical models by incorporating Social Network Exposure (SNE) and economic connectedness into the analysis of disease transmission and health behavior diffusion. Specifically, we examine how social structures, particularly cross-class friendships and strong clustering within social groups, impact the dynamics of both infectious diseases and chronic conditions.

Analysis 1: Economic Connectedness and Disease Spread investigates how social and economic ties influence the transmission of infectious diseases. Building on the Strength of Weak Ties Theory by Granovetter [5], this analysis introduces the Social Network Exposure (SNE) variable into the Susceptible-Infectious-Recovered (SIR) model [4]. By analyzing the impact of weak ties among disparate socioeconomic groups, this analysis elucidates how economic connectedness can facilitate or hinder the spread of diseases.

Analysis 2: Age-Structured Models for Understanding Infectious Disease Dynamics [6] emphasizes the importance of demographic factors in shaping infection rates and recovery patterns. This analysis segments populations into distinct age groups, reflecting variations in susceptibility and transmission rates. By utilizing age-structured models, it becomes evident that age-specific interactions significantly influence disease dynamics, thereby informing targeted interventions tailored to vulnerable age cohorts.

Analysis 3: The SLIAR Model [7] with Social Network Exposure (SNE) for Infectious Disease Dynamics extends the traditional SIR framework by incorporating both the latent and asymptomatic stages of infection. This model allows for a more nuanced understanding of how social ties and mobility affect disease progression. The inclusion of SNE within the SLIAR model captures the intricate interplay between social interactions and infectious disease dynamics, particularly within closely connected communities.

---

[*]4th year BS-MS Economics, Enrollment Number: 21322011

Analysis 4: Clustered Network Epidemics via Incorporating Social Interactions focuses on the impact of social clustering on disease spread [8]. By examining how infections propagate through clustered networks, this analysis reveals that social interactions significantly alter the trajectories of epidemics. Higher clustering coefficients lead to prolonged infection durations and sustained transmission within tight-knit groups, highlighting the role of community dynamics in shaping public health outcomes.

Together, these analyses provide a comprehensive framework for understanding the complexities of infectious disease transmission. By integrating economic connectedness, demographic factors, and social interactions into traditional epidemiological models, this report aims to inform public health policies that are responsive to the multifactorial nature of disease spread. Through this holistic approach, we can better address health disparities and improve health outcomes across diverse populations.

2 Summary of the Base Papers

Raj Chetty et al. have authored two seminal papers examining social capital and its role in economic mobility. The first paper, *Social Capital I: Measurement and Associations with Economic Mobility* [9], focuses on measuring social capital and its link to upward mobility, while the second, *Social Capital II: Determinants of Economic Connectedness* [10], investigates the factors that shape economic connections across different socioeconomic groups.

2.1 Social Capital I: Measurement and Associations with Economic Mobility

Defining and Measuring Social Capital

The authors describe social capital as the 'strength of an individual's connections within their social network and community' and explore three distinct types [9]:

- **Economic Connectedness (EC)**: This metric assesses the extent of the connection between individuals from low and high socioeconomic statuses (SES). It is determined by calculating the proportion of high-SES friends among low-SES individuals, adjusted relative to the total proportion of high-SES individuals in the population.
- **Network Cohesiveness**: This examines the structure of social networks, emphasizing the presence of cliques and the degree of mutual friendships. Metrics such as clustering coefficient, support ratio, and spectral homophily are used to quantify these aspects.
- **Civic Engagement**: It can be assessed through factors such as the rate of volunteering, engagement in community organizations, and the degree of trust within the community.

Key Findings

- **Geographical Variation**: There is considerable variation in all three social capital measures across different counties in the United States.
- **Economic Connectedness and Mobility**: The study finds a strong positive association between EC and upward income mobility. Children from low-income families who grow up in areas with higher EC tend to have higher incomes in adulthood. The study suggests that if children from low-SES families were raised in counties with EC compared to those experienced by the average child from high-SES parents, their future incomes would increase by an average of 20%. This factor stands out as one of the most significant predictors of upward economic mobility.
- **Other Social Capital Measures**: Network cohesiveness and civic engagement measures show weaker and less consistent correlations with economic mobility. For example, while civic engagement tends to be higher in rural areas, its correlation with economic mobility becomes weaker once the data is weighted to account for the greater concentration of low-income families in urban communities.

2.2 Social Capital II: Determinants of Economic Connectedness

This paper expands on the first paper by concentrating specifically on the factors that influence Economic Connectedness (EC), which is the form of social capital most closely associated with economic mobility. [10]

Decomposing Economic Connectedness

The authors distinguish between two key mechanisms that drive differences in EC:

- **Exposure**: The extent to which individuals from different SES backgrounds are present in the same social settings (e.g., schools, religious organizations, workplaces). The study reveals that around half of the disparity in Economic Connectedness (EC) between individuals from low and high-SES backgrounds can be attributed to differences in their exposure.
- **Friending Bias**: The tendency for individuals to form friendships with others from similar SES backgrounds, even when controlling for the level of exposure to people from different SES groups within a given setting. The study finds that the bias in forming friendships is more evident in larger, more diverse high schools.

Key Findings

- **Exposure and Bias Contribute Equally**: Both exposure and friending bias are key factors in influencing Economic Connectedness (EC), with each accounting for approximately half of the variation in the proportion of high-SES friends between low and high SES individuals.

- **Context Matters**: The extent of friending bias differs significantly across various social environments. For example, friending bias is generally less pronounced in religious organizations than in schools or workplaces.

Policy Implications

- **Integration for Low Bias**: In settings with low friending bias, increasing socioeconomic integration (exposure) can effectively boost EC.
- **Addressing Bias Directly**: In settings with high friending bias, interventions should focus on reducing bias within existing groups to promote cross-SES friendships.

3 Related Work

[11] analyzed the effect of reactive school closures on influenza spread by collecting contact diaries from 259 students and their households in Tomsk, Russia, modeling reduced transmission during school closures. Participants recorded daily contacts in three categories: (i) demographic information, (ii) contact patterns on regular school/workdays, and (iii) contact patterns during school closures. [12] assessed the limitations of the basic reproduction number R_0 in heterogeneous populations using multiplex networks and SIR models, proposing the time-varying $R(t)$ for a more accurate representation of epidemic dynamics. [13] applied a global stochastic epidemic model to investigate the transmission of Zika throughout the Americas, showing that mosquito dynamics, mobility, and seasonal factors influenced the outbreak's spatial and temporal patterns. The study utilizes the Global Epidemic and Mobility Model (GLEAM), a stochastic, agent-based framework that uses high-resolution data for various population layers, including households, schools, workplaces, and communities. [14] studied SARS-CoV-2 transmission in Ethiopia, finding higher infection rates in urban areas but more severe cases in rural regions, linking socio-demographic factors to disease burden in sub-Saharan Africa. [15] explored how opinion dynamics on disease severity impact infection spread in networks using a discrete-time SIS model, highlighting the role of social factors in epidemiological models.

[8] examined epidemic spread in clustered networks, finding that clustering increases the epidemic threshold by reducing component sizes and transmission pathways compared to unclustered networks. [16] proposed a method for estimating spatially explicit reproduction numbers $R_k(t)$ by incorporating mobility dynamics, improving epidemic modeling with real COVID-19 data from Italy. [17] revealed that SARS-CoV-2 spread undetected in early 2020 due to limited testing, with international travel driving early introductions. By March, only 1–3% of infections were detected in the US and Europe. [18] found demographic biases in the Influenzanet surveillance network, with underrepresentation of certain age groups, smokers, and diabetics impacting the generalizability of influenza surveillance data. [19] studied chikungunya transmission in Bangladesh, highlighting the significant role of household-level social interactions, with gender differences and proximity being key factors in infection risk. The study integrates detailed epidemiological data with statistical modeling to assess various risk factors associated with disease transmission.

4 Data and Measures

4.1 Data

The dataset in the base papers [9] and [10] focuses on friendship links among Facebook users between the ages of 25 and 44 living in the United States. The dataset encompasses 21 billion friendship pairs, representing a significant portion of the U.S. population within that age range. The authors use this data to assess different aspects of social capital, such as economic connectedness, network cohesion, and civic participation. The study focuses on this age group because their Facebook usage rate exceeds 80%, surpassing other age groups, and aligns with readily available demographic data from the American Community Survey (ACS), allowing for comparisons with the broader population.

User Information: In addition to friendship links, the researchers extract data from users' Facebook profiles, which includes self-reported information on several variables such as age, location (ZIP code and county), language, relationship status, educational attainment, gender, college attended, donations made, phone model price, usage of Facebook on different devices (mobile vs. internet), familial ties, mobile carrier, and group memberships. The researchers integrate these variables to create a socioeconomic status (SES) measure for each user.

Parental SES and Group-Level Data: To estimate family SES during childhood, the authors connect individuals in their primary dataset to their parents by utilizing self-reported family relationships, a hashed version of users' last names, and publicly accessible posts and significant life events. In order to examine the social context of connections, the researchers classify friendships into various social settings based on common affiliations, such as high schools, universities, workplaces, neighborhoods (ZIP codes), religious groups, and recreational organizations.

External Datasets: In addition to the Facebook data, the studies incorporate a range of publicly accessible datasets, such as the American Community Survey (ACS), which is used to benchmark the sample and gather information on median incomes by block group, racial and ethnic composition, and other community features. The Opportunity Atlas offers data on intergenerational income mobility, high school graduation rates, and teenage birth rates. School-level data is sourced from the National Center for Education Statistics (NCES) and the Civil Rights Data Collection (CRDC), while the Integrated Postsecondary Education Data System (IPEDS) provides college-level data. Additionally, Census Data is used to identify the number of individuals with below-national-median parental income, which informs the weighting of correlations and regressions.

4.2 Statistical Analysis

This report primarily focuses on theoretical frameworks to understand the dynamics of infectious disease transmission. Simulations have been conducted to illustrate the model outcomes and validate the theoretical constructs. The simulations were implemented using Python, with the accompanying code provided in the attached folder. This computational approach allows for the exploration of various scenarios and the assessment of the impact of different variables on disease dynamics. The authors calculated the exposure and clustering coefficient, which are utilized in this paper.

4.3 Measures

- **Infection Population:** Represented by the number of individuals actively infected at a specific point in time.
- **Susceptible Population:** Refers to the individuals who are not currently infected but are vulnerable to contracting the disease.
- **Latent Period:** The period in which individuals are infected but have not yet become contagious. This period is crucial for understanding the dynamics of disease spread.
- **Social Network Exposure (SNE):** Quantified based on the density of social ties within communities, measured through network analysis techniques. The SNE varies between 0 and 1, with higher values signifying stronger social and economic connectedness.
- **Clustering Coefficient:** Calculated to assess the degree of interconnectedness among individuals within social networks. Similar to the SNE, the clustering coefficient spans from 0 to 1, where higher values indicate more closely-knit social groups.

5 Theoretical Frameworks

5.1 Analysis 1: Economic Connectedness and Disease Spread

Existing Theory: The Strength of Weak Ties

Strength of Weak Ties theory (1973) by Granovetter [5] is foundational in social network research, positing that weak ties between disparate groups are critical for the diffusion of information. These weak ties, typically characterized as connections between individuals from different socioeconomic backgrounds, provide a bridge for sharing new ideas, knowledge, and opportunities that would otherwise remain isolated within homogeneous groups.

In the context of disease transmission, this theory can be adapted to highlight how cross-class friendships (weak ties) facilitate faster and broader disease spread by connecting different socioeconomic groups. When applied to an epidemiological framework, weak ties serve as conduits through which infections can travel across the otherwise disconnected clusters of social groups, thereby increasing the overall exposure in a population. Conversely, strong homophily (friending bias), where people preferentially form relationships within their own social and economic class, limits this cross-group transmission and results in more contained outbreaks.

Extension: Social Network Exposure (SNE) in the SIR Model

To examine how economic connectedness influences disease transmission, we introduce a Social Network Exposure (SNE) variable into the classical SIR (Susceptible-Infectious-Recovered) model. This variable measures the level of interaction between individuals from different socioeconomic backgrounds, capturing the economic connectedness across social classes. By modifying the infection rate based on this level of social interaction, we are able to model how cross-class friendships influence the overall dynamics of infectious disease spread.

The SIR model [4] describes the evolution of three key population groups:

- Susceptible (S): Individuals who are at risk of contracting the disease.
- Infected (I): Individuals who are currently infected and capable of spreading the disease.
- Recovered (R): Individuals who have recovered from the disease and are immune to reinfection.

The higher the **SNE** value, the more frequent the interactions between different socioeconomic groups, leading to increased disease spread. Conversely, lower **SNE** indicates higher friending bias, where cross-class interactions are limited, resulting in fewer cross-group infections. The equations are adjusted as follows:

1. **Modified Susceptible Population**:

$$\frac{ds(t)}{dt} = -\beta \cdot k \cdot \text{SNE} \cdot i(t) \cdot [1 - r(t) - i(t)]$$

Where:

- $s(t)$: The fraction of the population that is susceptible at time t.
- β: The infection rate, representing the likelihood of disease transmission per contact.
- k: The average number of interactions an individual has.
- $i(t)$: The fraction of the population that is infected at time t.
- $r(t)$: The fraction of the population that has recovered from the disease.

The SNE modifies the infection rate based on the degree of cross-class interactions, increasing transmission as economic connectedness rises.

2. **Modified Infected Population**:

$$\frac{di(t)}{dt} = \beta \cdot k \cdot \text{SNE} \cdot i(t) \cdot [1 - r(t) - i(t)] - \mu \cdot i(t)$$

Where:

- μ: Recovery rate, indicating how quickly infected individuals recover.

The infection term now reflects the influence of SNE, where higher cross-class interactions boost the probability of infected individuals contracting the disease [20].

   3. **Recovered Population**:

$$\frac{dr(t)}{dt} = \mu \cdot i(t)$$

Where:

- $r(t)$: Fraction of the population that has recovered at time t.

5.2 Analysis 2: Age-Structured Models for Understanding Infectious Disease Dynamics

Age-structured models [6] are essential in infectious disease research as they consider the variation in transmission rates, susceptibility, and recovery across different age groups. These models are vital for understanding how diseases spread in populations with heterogeneous contact patterns and immune responses. By segmenting populations by age, these models provide a detailed insight into disease behavior and facilitate the creation of more focused intervention strategies.

   Key Components of Age-Structured Models:
   1. **Population Segmentation:**
The population is divided into distinct age groups, or cohorts, each with unique characteristics related to:

- **Infection Rates:** Younger individuals may have higher rates of contact, such as through schools, and thus higher transmission potential, while older individuals may be more vulnerable to severe outcomes.
- **Recovery Rates:** Recovery can vary with age due to immune system differences, with younger individuals often recovering faster than older ones.
- **Contact Patterns:** The model accounts for how different age groups interact with one another, with children, for example, having more frequent close contact with peers, while older individuals may have fewer but more high-risk interactions.

   2. **Dynamic Interactions:**
Age-structured models simulate the interactions between these age groups, allowing researchers to study how a disease spreads through the different cohorts over time based on varying infection and recovery rates.

   Basic Structure of Age-Structured Models
The population is divided into several age classes, denoted as $A_1, A_2, \ldots, A_n$, where each class A_i represents a specific age range (e.g., 0-5 years, 6-15 years, 16-59 years, 60+ years). The model typically uses a compartmental approach for each age group, with individuals classified into infectious (I), susceptible (S), and recovered (R) compartments.

   Equations of Age-Structured Models
   1. **Susceptible Population:**

$$\frac{dS_i}{dt} = -\beta_i S_i \sum_j \left(\frac{I_j}{N_j} \right) \cdot C_{ij}$$

The variable S_i represents the count of individuals in age group A_i who are susceptible to the disease. β_i denotes the transmission rate specific to age group A_i. I_j refers to the count of individuals in age group A_j who are currently infected. The element C_{ij} of the contact matrix quantifies the interaction rates between age groups A_i and A_j.

$$C_{ij} = SNE_{ij}$$

Where: - SNE_{ij}: The Social Network Exposure factor that quantifies the level of cross-class interaction between age groups A_i and A_j. Higher values indicate more interaction and greater economic connectedness.

   2. **Infectious Population:**

$$\frac{dI_i}{dt} = \beta_i S_i \sum_j \left(\frac{I_j}{N_j} \right) \cdot C_{ij} - \gamma_i I_i$$

The variable I_i represents the number of individuals in age group A_i who are infectious. γ_i denotes the recovery rate for age group A_i.

   3. **Recovered Population:**

$$\frac{dR_i}{dt} = \gamma_i I_i$$

- R_i: The count of individuals in age group A_i who have recovered.

   These equations describe the transitions between susceptible, infected, and recovered states within each age group, while the contact matrix C captures the interactions between different age groups, influencing the spread of the disease.

   Matrix Representation of the Model
In matrix form, the system of equations can be represented as:

$$\mathbf{X} = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_n \\ I_1 \\ I_2 \\ \vdots \\ I_n \\ R_1 \\ R_2 \\ \vdots \\ R_n \end{bmatrix} \quad \text{and} \quad \mathbf{C} = \begin{bmatrix} C_{11} & C_{12} & \ldots & C_{1n} \\ C_{21} & C_{22} & \ldots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \ldots & C_{nn} \end{bmatrix}$$

The contact matrix $\mathbf{C}$ represents the rates of interaction between various age groups, influencing the transmission of the disease.

The system of equations can be summarized as follows:

$$\frac{d\mathbf{X}}{dt} = \mathbf{A}(\mathbf{X}) - \mathbf{B}(\mathbf{X})$$

Where $\mathbf{A}$ represents the infection dynamics, and $\mathbf{B}$ represents the recovery dynamics.

5.3 Analysis 3: SLIAR Model with Social Network Exposure (SNE) for Infectious Disease Dynamics

The SLIAR model [7] is an extension of the traditional SIR (Susceptible-Infectious-Recovered) framework that incorporates additional compartments to account for the latent and asymptomatic stages of infection. This model allows for a more nuanced understanding of how infectious diseases spread, particularly in populations where asymptomatic individuals play a significant role.

1. Compartment Definitions
The SLIAR model includes the following compartments:

- **S (Susceptible):** Individuals who are not infected and can become infected.
- **L (Latent):** Individuals who have contracted the disease but are not yet able to transmit it. They cannot transmit the disease to others during this latent period but will eventually progress to the infectious state.
- **I (Infectious Symptomatic):** Individuals who are infected and exhibit symptoms. These individuals can spread the disease to susceptible individuals.
- **A (Infectious Asymptomatic):** Individuals who are infected but do not show symptoms. They can still transmit the disease, often without being aware of their infection.
- **R (Removed):** Individuals who have either recovered from the disease and developed immunity or who have passed away. They can no longer spread the disease.

2. Modification of the SLIAR Model with Social Network Exposure (SNE)
To incorporate Social Network Exposure (SNE) into the SLIAR model, we modify the equations to account for how social interactions influence disease transmission dynamics. The SNE variable reflects the degree of cross-class interactions within the community, affecting the infection rates.

Modified SLIAR Model Equations:

1. **Susceptible Population:**

$$\frac{dS}{dt} = -\beta S \left(I + \delta A \right) \cdot SNE, \quad S(0) = S_0 > 0$$

SNE represents a factor that quantifies social network exposure, with higher values indicating more interaction and thus a greater chance of infection [20]. The parameter β refers to the transmission rate, while I denotes infectious symptomatic individuals. The term A represents infectious asymptomatic individuals, and δ is a factor that accounts for the increased risk of infection posed by asymptomatic individuals.

2. **Latent Population:**

$$\frac{dL}{dt} = \beta S \left(I + \delta A \right) \cdot SNE - \kappa L, \quad L(0) = L_0 \geq 0$$

The latent individuals' infection dynamics are also influenced by the SNE factor. The parameter κ represents the rate at which individuals transition from the latent phase to the infectious phase.

3. **Infectious Symptomatic Population:**

$$\frac{dI}{dt} = p\kappa L - \alpha I, \quad I(0) = I_0 \geq 0$$

6

The parameter p represents the proportion of latent individuals who become symptomatic, while α is the recovery rate for symptomatic individuals. This equation remains unchanged as the progression of symptomatic individuals is not directly influenced by the SNE.

4. **Infectious Asymptomatic Population:**

$$\frac{dA}{dt} = (1-p)\kappa L - \eta A, \quad A(0) = A_0 \geq 0$$

The parameter η represents the recovery rate for asymptomatic individuals. This equation also remains unchanged.

5. **Removed Population:**

$$\frac{dR}{dt} = f\alpha I + \eta A, \quad R(0) = R_0 \geq 0$$

The parameter f denotes the proportion of symptomatic individuals who transition to the removed class due to recovery or death. This equation also remains unchanged.

6. **Total Population:**

$$N(t) = S(t) + L(t) + I(t) + A(t) + R(t)$$

The inclusion of SNE modifies the transmission dynamics for both the Susceptible and Latent compartments. By capturing the effect of varying degrees of social interactions and economic connectedness, the model better reflects real-world disease dynamics.

5.4 Analysis 4: Clustered Network Epidemics via Incorporating Social Interactions

This approach is particularly relevant in studying how social interactions influence the spread of infectious diseases among closely-knit groups [8]. Here, the clustering coefficient is used.

1. Infection Probability in a Triangle (Social Group)

The model treats nodes u, v, and w as members of a triangular social group, representing tight-knit relationships such as family, friends, or coworkers. This configuration captures how infections spread among interconnected individuals.

Infection Probabilities: - **Both Neighbors Infected**:

$$P(\text{infect both}) = C \cdot (3T^2 - 2T^3)$$

- **Exactly One Neighbor Infected**:

$$P(\text{infect one}) = 2T(1-T)^2$$

- **Neither Neighbor Infected**:

$$P(\text{infect none}) = (1-T)^2$$

Here, T is the transmissibility, reflecting the probability of infection occurring along an edge. The clustering coefficient C enhances the likelihood of simultaneous infections within these close-knit groups.

Expected Number of Infections: The expected number of infections per triangle can be computed as follows:

$$E[\text{Infections}] = 2T(1-T)^2 + 2C(3T^2 - 2T^3) = 2T + 4T^2 - 4T^3 + 6CT^2 - 4CT^3$$

This equation illustrates the dependence of infection spread on both transmissibility T and the clustering coefficient C.

2. Next-Generation Matrix for Social Network Exposure (SNE)

To capture the transmission dynamics through different pathways—independent edges and triangle edges—a next-generation matrix is defined:

$$\begin{bmatrix} n_I(s+1) \\ n_\Delta(s+1) \end{bmatrix} = \begin{bmatrix} c_{II} & c_{I\Delta} \\ c_{\Delta I} & c_{\Delta\Delta} \end{bmatrix} \begin{bmatrix} n_I(s) \\ n_\Delta(s) \end{bmatrix}$$

Where: - $n_I(s)$: Number of nodes infected through independent edges. - $n_\Delta(s)$: Number of nodes infected through triangle edges.

Calculating the Matrix Elements: 1. Infection along Independent Edges:

$$c_{II} = T \cdot \frac{\langle K_I^2 - K_I \rangle}{\langle K_I \rangle}$$

2. Infection along Triangle Edges from Independent Edges:

$$c_{I\Delta} = (2T + 4T^2 - 4T^3 + 6CT^2 - 4CT^3) \cdot \frac{\langle K_I K_\Delta \rangle}{\langle K_I \rangle}$$

3. Infection along Independent Edges from Triangle Edges:

$$c_{\Delta I} = T \cdot \frac{\langle K_I K_\Delta \rangle}{\langle K_\Delta \rangle}$$

7

4. Infection along Triangle Edges:

$$c_{\Delta\Delta} = (2T + 4T^2 - 4T^3 + 6CT^2 - 4CT^3) \cdot \frac{\langle K_\Delta^2 - K_\Delta \rangle}{\langle K_\Delta \rangle}$$

These matrix elements capture the pathways through which infections spread and highlight the importance of both individual and clustered interactions in the overall transmission dynamics.

3. Basic Reproduction Number and Critical Transmissibility

The basic reproduction number, R_0, is a crucial epidemiological metric that indicates the potential for disease spread. It is determined from the dominant eigenvalue of the next-generation matrix, which models the transmission dynamics of the infection. If R_0 exceeds 1, it suggests that the infection has the potential to spread through the population. Conversely, when R_0 is less than 1, it implies that the disease will eventually fade out.

6 Results

6.1 Analysis 1: Economic Connectedness and Disease Spread

The simulation of the SIR model with Social Network Exposure (SNE), as shown in Figure 1, was conducted using Python with a total population of 1,000 individuals. Initially, there was one infected individual and 0 recovered individuals, resulting in 999 susceptible individuals. The contact rate (β) was set at 0.2, while the mean recovery rate (γ) was $1/10$, indicating an average recovery time of 10 days. The SNE factor was fixed at 1, reflecting maximum social network exposure. The simulation spanned 160 days, with time points generated at 1-day intervals using 'numpy.linspace'. The SIR model equations were modified to incorporate SNE and the resulting differential equations were solved using the 'odeint' function from the 'scipy.integrate' library to assess the impact of social interactions on disease dynamics. Figure 2 is obtained by changing SNE.
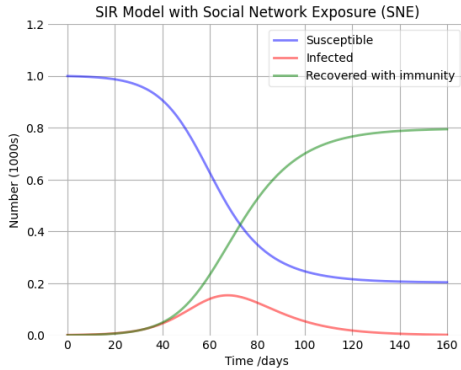

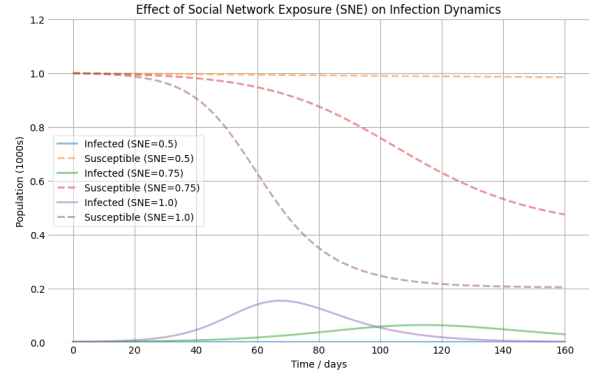Figure 1: SIR Model with Social Network Exposure (SNE) =1

Figure 2: Effect of Social Network Exposure (SNE) on Infection Dynamics

In regions with high economic connectedness (low friending bias), the disease spreads rapidly across socioeconomic groups, leading to an earlier peak in infections and a faster decline in susceptible individuals, as shown in Figure 2. This accelerated transmission strains healthcare resources sooner, requiring prompt interventions. Conversely, in areas with high friending bias (low economic connectedness), outbreaks are more localized within socioeconomic groups, resulting in slower, stratified transmission. This delays the infection peak and causes a more gradual decline in susceptible individuals.

6.2 Analysis 2: Age-Structured Models for Understanding Infectious Disease Dynamics

The age-structured model simulation, as shown in Figure 3, was executed using Python with a total population of 1,000 individuals, evenly divided between two age groups: young and old. Initial conditions included one infected individual in each age group, with zero recovered individuals. The susceptible populations were calculated as follows: the young susceptible population was set to $S_y = 499$ (i.e., half of the population minus the infected), and the old susceptible population was set to $S_o = 499$. Contact rates varied between age groups, with the transmission rate for the young set at 0.3 and for the old at 0.1. The mean recovery rate for both groups was defined as $\frac{1}{10}$ and $\frac{1}{10}$ respectively, indicating an average recovery period of 10 days. The Social Network Exposure (SNE) factor was fixed at 1, signifying maximum exposure. The simulation ran for 160 days, using time points generated at 1-day intervals with 'numpy.linspace'. The resulting differential equations were solved using the 'odeint' function from the 'scipy.integrate' library, allowing for the assessment of disease dynamics across age groups and the influence of social interactions. Figure 4 is obtained by changing SNE.

Higher Social Network Exposure (SNE) increases cross-class interactions, leading to faster disease transmission and earlier infection peaks across age groups, as shown in Figure 4. This acceleration of transmission is particularly evident when younger individuals, often in schools or community settings, have more frequent contact with older adults, heightening the risk of rapid disease spread. As a result, public health systems must prepare for surges in cases earlier than anticipated, requiring swift response measures. Additionally, changes in SNE can highlight vulnerabilities in certain age groups, especially older adults who
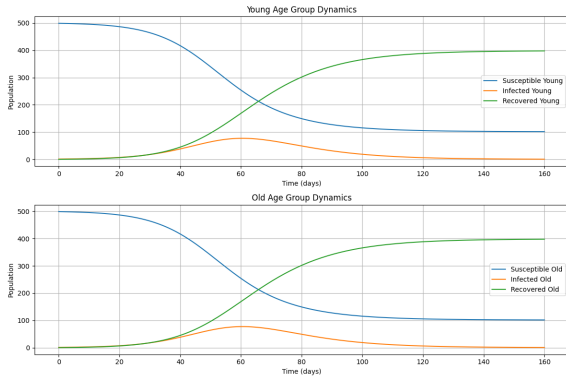
Figure 3: SIR Model with Young and Old Age Group Dynamics with Social Network Exposure (SNE) =1
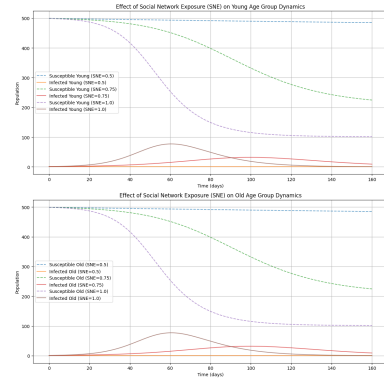

Figure 4: Effect of Social Network Exposure (SNE) on Infection Dynamics in Young and Old Age Group SIR Model

may be exposed to younger, asymptomatic carriers. The distribution of risk shifts accordingly, necessitating targeted monitoring and intervention. Conversely, lower SNE limits interactions across socioeconomic groups, leading to more localized outbreaks. While this can make containment strategies more straightforward, it may also strain local healthcare resources if not managed effectively.

6.3 Analysis 3: SLIAR Model with Social Network Exposure (SNE) for Infectious Disease Dynamics

The simulation of the SLIAR model with Social Network Exposure (SNE), as shown in Figure 5, was conducted using Python, employing a transmission rate of 0.5 and an infectivity reduction rate for asymptomatic individuals set at 0.3. The transition rate from latent to infected was defined as 0.1, while the recovery rate for symptomatic infections was set at 0.2. The proportion of latent individuals progressing to symptomatic infections was established at 0.6, with a recovery rate for symptomatic individuals fixed at 0.95. Asymptomatic individuals had a recovery rate of 0.1. The SNE factor was fixed at 1.0, indicating full exposure. Initial conditions included 99% of the population as susceptible, 1% as latent, and none as symptomatic, asymptomatic, or recovered. The simulation spanned 160 days, with time points generated at 1-day intervals, and the differential equations were solved using the 'odeint' function from the 'scipy.integrate' library. Figure 6 is obtained by changing SNE.
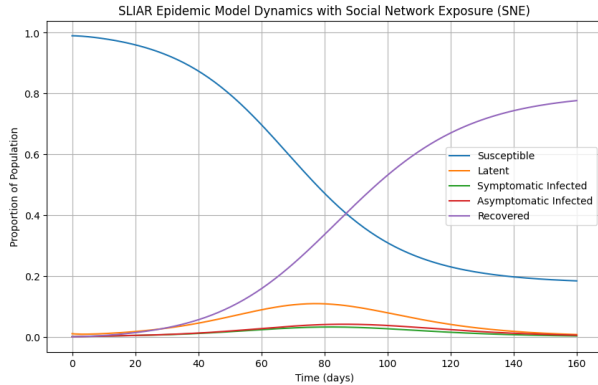

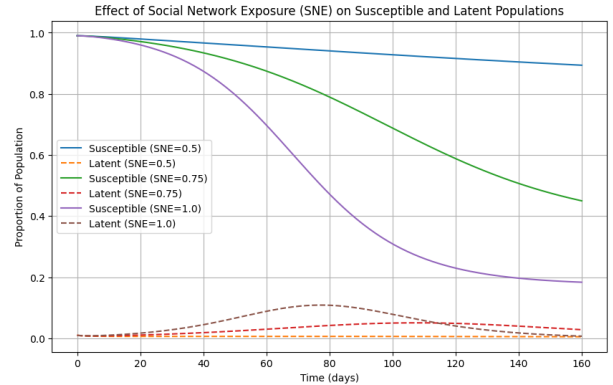Figure 5: SLIAR Epidemic Model Dynamics with Social Network Exposure (SNE) = 1

Figure 6: Effect of Social Network Exposure (SNE) on SLIAR Epidemic Model Dynamics

Higher Social Network Exposure (SNE) leads to more frequent interactions across socioeconomic groups, resulting in increased infection rates as individuals rapidly move from the susceptible (S) to latent (L) and infectious (I, A) compartments, as shown in Figure 6. This accelerates disease transmission, particularly among asymptomatic carriers, and drives an earlier peak in infections, putting additional pressure on healthcare systems. Shortened latent and infectious periods further speed up the spread, while asymptomatic individuals contribute significantly to undetected transmission. Different age groups face varying vulnerabilities, with younger populations becoming key transmitters. High SNE also leads to localized outbreaks across socioeconomic groups, complicating containment efforts, whereas low SNE confines outbreaks within specific classes. This stratification may result in disparities in infection and recovery outcomes.

6.4 Analysis 4: Clustered Network Epidemics via incorporating social interactions

For Figure 7, a clustered network consisting of 1,000 nodes was generated, where the probability of forming triangle connections was set at 0.3. The degree distribution for independent edges was modeled using a Poisson distribution with a mean of 2, while triangle edges had a mean of 1. The probability of infection was set at 0.1, while the recovery probability was 0.05. One node was randomly selected as the initial infected individual, and the simulation was run for a maximum of 100-time steps,

plotting the number of infected, susceptible, and recovered individuals over time. For Figure 8, the same parameters were used, but the clustering coefficient of the network was explicitly incorporated into the infection dynamics. The clustering coefficient was calculated based on the average connections in the network, influencing the likelihood of infection spread between tightly connected groups.
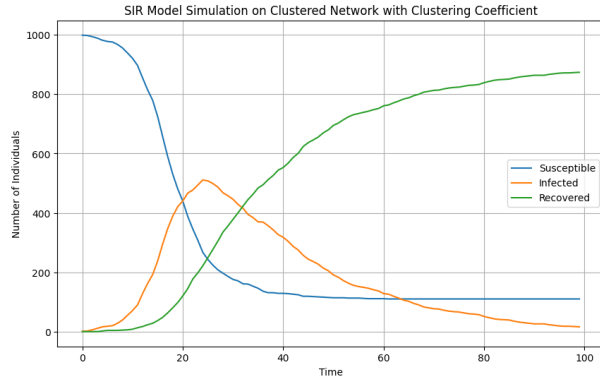

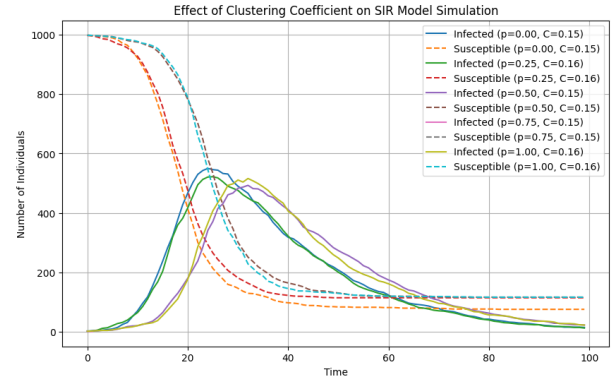Figure 7: SIR Model Simulation on Clustered Network with Clustering Coefficient

Figure 8: Effect of Clustering Coefficient on SIR Model Simulation of Clustered Network

The Clustering Effect on Infection Spread shows that at $p = 0.00$, infections quickly die out. As p increases (e.g., 0.25, 0.50, 0.75, 1.00), infections peak before declining, indicating that higher infection probabilities lead to faster, more widespread transmission (Figure 8). A larger clustering coefficient ($C = 0.16$) prolongs the infection duration due to more efficient spread within closely connected social groups, while lower clustering results in quicker but shorter epidemics. Social Interactions play a crucial role—higher clustering keeps infections active longer, while higher infection probabilities p increase the epidemic's peak magnitude.

7 Public Health Implications

In regions with high economic connectedness, broad public health interventions, such as widespread vaccination and social distancing, are essential to mitigate rapid disease transmission across all socioeconomic groups. In contrast, areas with low economic connectedness may benefit more from targeted interventions focused on specific communities, such as low-SES populations where diseases may be more contained.

Targeted Interventions: Public health efforts should prioritize vaccination for high-risk groups identified through Social Network Exposure (SNE) analysis, particularly younger populations acting as key transmitters. Tailored health education campaigns should raise awareness about inter-age group risks and promote preventive measures in high-SNE settings.

Flexible Response Strategies: Public health responses must be adaptive to changes in SNE, implementing temporary restrictions or enhanced measures during community events. Real-time monitoring of SNE can allow for more proactive strategies, enabling health officials to respond dynamically to evolving conditions.

Community Engagement: Encouraging cross-class interactions can strengthen community resilience to outbreaks. Public policies should support initiatives that promote inter-socioeconomic interactions alongside providing essential support services like transportation for older adults to healthcare facilities.

Resource Allocation: Healthcare resources should be allocated based on age-specific risk and SNE patterns. High-SNE areas may require increased medical support due to faster transmission rates, while strategic healthcare infrastructure planning should account for potential spikes in infections among populations frequently interacting across age groups.

8 Conclusion

The research presented in this paper highlights the critical role of social network exposure (SNE), connectedness, clustering co-efficient, and age-structured dynamics in comprehending and controlling the transmission of infectious diseases. By extending traditional models such as SIR and SLIAR to incorporate SNE, this study demonstrates how cross-class and inter-age interactions accelerate disease transmission, leading to different public health outcomes in various socioeconomic contexts. The simulations of clustered network epidemics highlight the impact of social connections and clustering on disease transmission, providing insights into how closely connected groups may sustain transmission longer.

The findings underscore the need for adaptive and targeted public health strategies, with interventions tailored to specific population dynamics and network structures. These insights have significant implications for vaccine distribution, resource allocation, and community engagement initiatives. Monitoring SNE and integrating it into public health planning can lead to more effective, proactive measures in both highly connected and more isolated regions.

Future research should focus on refining these models with real-world data to validate the findings and explore additional factors, such as mobility patterns and health behaviors, that can influence disease dynamics. Ultimately, a deeper understanding of how social structures impact epidemic spread can inform more efficient and equitable public health policies.

9 Data and Code Availability

All data, code, and materials used in this study are available in a GitHub repository for transparency and reproducibility: https://github.com/divya830/Modeling_Disease_Spread_and_Health_Behavior_Diffusion_through_Social_Network_Exposure. This repository provides access to the full set of resources used in this research.

References

[1] L. F. Berkman and T. Glass, "Social integration, social networks, social support, and health," in *Social Epidemiology*. Oxford University Press, 2000.

[2] M. Marmot, "Social determinants of health inequalities," *The Lancet*, vol. 365, no. 9464, pp. 1099–1104, 2005. [Online]. Available: https://doi.org/10.1016/S0140-6736(05)71146-6

[3] A. V. Diez Roux, "Investigating neighborhood and area effects on health," *American Journal of Public Health*, vol. 91, no. 11, pp. 1783–1789, 2001. [Online]. Available: https://doi.org/10.2105/AJPH.91.11.1783

[4] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 115, no. 772, pp. 700–721, 1927. [Online]. Available: https://doi.org/10.1098/rspa.1927.0118

[5] M. S. Granovetter, "The strength of weak ties," *American Journal of Sociology*, vol. 78, no. 6, pp. 1360–1380, 1973. [Online]. Available: https://doi.org/10.1086/225469

[6] O. Diekmann, H. A. P. Heesterbeek, and T. Britton, *Mathematical Tools for Understanding Infectious Disease Dynamics*. Princeton University Press, 2013.

[7] M. Ajelli and A. Vespignani, "The sliar model: a framework for modeling the dynamics of infectious diseases with latent and asymptomatic infections," *Journal of Theoretical Biology*, vol. 266, no. 1, pp. 121–130, 2010.

[8] J. C. Miller, "Percolation and epidemics in random clustered networks," *Physical Review E*, vol. 80, no. 2, p. 020901, 2009. [Online]. Available: https://doi.org/10.1103/PhysRevE.80.020901

[9] R. Chetty, M. O. Jackson, T. Kuchler, J. Stroebel, N. Hendren, R. B. Fluegge, S. Gong, F. Gonzalez, A. Grondin, M. Jacob, D. Johnston, M. Koenen, E. Laguna-Muggenburg, F. Mudekereza, T. Rutter, N. Thor, W. Townsend, R. Zhang, M. Bailey, P. Barberá, M. Bhole, and N. Wernerfelt, "Social capital i: measurement and associations with economic mobility," *Nature*, vol. 608, 2022. [Online]. Available: https://doi.org/10.1038/s41586-022-04996-4

[10] R. Chetty, M. O. Jackson, T. Kuchler, J. Stroebel, N. Hendren, R. B. Fluegge, S. Gong, F. Gonzalez, A. Grondin, M. Jacob, D. Johnston, M. Koenen, E. Laguna-Muggenburg, F. Mudekereza, T. Rutter, N. Thor, W. Townsend, R. Zhang, M. Bailey, P. Barber, M. Bhole, and N. Wernerfelt, "Social capital ii: determinants of economic connectedness," *Nature*, vol. 608, 2022. [Online]. Available: https://doi.org/10.1038/s41586-022-04997-3

[11] M. Litvinova, Q.-H. Liu, E. S. Kulikov, and M. Ajelli, "Reactive school closure weakens the network of social interactions and reduces the spread of influenza," *Proceedings of the National Academy of Sciences*, vol. 116, no. 27, pp. 13 174–13 181, 2019. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.1821298116

[12] Q.-H. Liu, M. Ajelli, A. Aleta, S. Merler, Y. Moreno, and A. Vespignani, "Measurability of the epidemic reproduction number in data-driven contact networks," *Proceedings of the National Academy of Sciences*, vol. 115, no. 50, pp. 12 680–12 685, 2018. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.1811115115

[13] Q. Zhang, K. Sun, M. Chinazzi, A. P. y Piontti, N. E. Dean, D. P. Rojas, S. Merler, D. Mistry, P. Poletti, L. Rossi, M. Bray, M. E. Halloran, I. M. Longini, and A. Vespignani, "Spread of zika virus in the americas," *Proceedings of the National Academy of Sciences*, vol. 114, no. 22, pp. E4334–E4343, 2017. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.1620161114

[14] F. Trentini, G. Guzzetta, M. Galli *et al.*, "Modeling the interplay between demography, social contact patterns, and sars-cov-2 transmission in the south west shewa zone of oromia region, ethiopia," *BMC Medicine*, vol. 19, no. 1, p. 89, 2021. [Online]. Available: https://doi.org/10.1186/s12916-021-01967-w

[15] W. Xuan, R. Ren, and J. Liu, "On a discrete-time network sis model with opinion dynamics," *IFAC-PapersOnLine*, vol. 54, no. 13, pp. 618–623, 2021. [Online]. Available: https://doi.org/10.1016/j.ifacol.2021.04.176

[16] C. Trevisin, E. Bertuzzo, D. Pasetto, L. Mari, S. Miccoli, R. Casagrandi, M. Gatto, and A. Rinaldo, "Spatially explicit effective reproduction numbers from incidence and mobility data," *Proceedings of the National Academy of Sciences*, vol. 120, no. 20, p. e2219816120, 2023. [Online]. Available: https://doi.org/10.1073/pnas.2219816120

[17] J. T. Davis, M. Chinazzi, N. Perra, K. Mu, A. Pastore y Piontti, M. Ajelli, N. E. Dean, C. Gioannini, M. Litvinova, S. Merler *et al.*, "Cryptic transmission of sars-cov-2 and the first covid-19 wave in europe and the united states," *medRxiv*, 2021, preprint. [Online]. Available: https://doi.org/10.1101/2021.03.24.21254199

[18] P. Cantarelli, M. Debin, C. Turbelin *et al.*, "The representativeness of a european multi-center network for influenza-like-illness participatory surveillance," *BMC Public Health*, vol. 14, no. 1, p. 984, 2014. [Online]. Available: https://doi.org/10.1186/1471-2458-14-984

[19] H. Salje, J. Lessler, K. K. Paul, and S. Cauchemez, "How social structures, space, and behaviors shape the spread of infectious diseases using chikungunya as a case study," *Proceedings of the National Academy of Sciences*, vol. 113, no. 47, pp. 13 420–13 425, 2016. [Online]. Available: https://doi.org/10.1073/pnas.1611391113

[20] H. Nunner, V. Buskens, and M. Kretzschmar, "A model for the co-evolution of dynamic social networks and infectious disease dynamics," *Computational Social Networks*, vol. 8, p. 19, 2021. [Online]. Available: https://doi.org/10.1186/s40649-021-00098-9