Economics of Prison: Modeling the Dynamics of Recidivism

Alejandra Lopez[1], Naomi Moreira[2], Ariana Rivera[3], Bechir Amdouni[1], Baltazar Espinoza[1] and Christopher M Kribs[4]

[1]Arizona State University
[2]St. Joseph's College
[3]Yachay Tech University
[4]University of Texas at Arlington

Abstract

In Arizona each prisoner costs the state an average of $25,397 per year, the approximate cost of attending Arizona State University [21]. Based on the current population of inmates this adds up to over 1 billion dollars annually. This figure is 5 times more than is spent on public assistance and about 70 percent of what is spent on transportation in the entire state. In addition to that, 40 percent of those who leave the prison return which further increases the costs on the state. In an effort to decrease costs the government of Arizona hopes to implement programs into their prison system in order to lower the recidivism rate and decrease costs. Multiple studies have shown recidivism is reduced when education and transition programs are incorporated. Currently, Arizona funds most of the GED program, excluding testing, while an inmate is incarcerated. Unfortunately, this education does not continue after the inmate is released. Meanwhile, other states have successfully incorporated education in order to reduce recidivism. In an effort to analyze recidivism in Arizona we have developed and analyzed a data-driven mathematical model that captures the dynamics of prisoners while in and out of prison based on their education status. This model, a system of differential equations, helped to estimate the cost associated with different educational programs in and outside of prison to the cost of recidivism. As a result we were able to study the economic impact of implementing these transition programs which we proved to be cost efficient. We found that the transition programs would eventually pay for themselves as a higher proportion of inmates enroll in the program. We were also able to show that it was possible to completely eliminate recidivism as the length of the program increased and enough inmates enrolled in the transition program after being released.

1 Introduction

In 1984, the United States entered the "tough on crime" era which resulted in mandatory minimum sentences for drug offenses [17]. As a result of this measure, incarceration rates increased rapidly over a short period of time. Presently, the United States has the highest incarceration rate per year in the world with over 2 million people imprisoned [23]. That is roughly 6 times higher than China's incarceration rate per year, a country with a population 4 times larger than the population of the United States. [3] Incarceration and high recidivism rates come with high economic and social costs. Consequently, various states have invested a great amount of effort into minimizing the economic impact on their budget. We consider how effective educational correction programs are, specifically the General Education Diploma (GED) inside of the state prisons and the reentry programs outside that parolees have access to in Arizona.

Over the last thirty years, the inmate population in Arizona has increased from over ten thousand inmates to more than forty two thousand inmates, an increase of 293%. To accommodate the increased population of inmates, the corrections' budget also increased dramatically over the same time period going from $211.5 million to $1.1 billion, an increase of 419% [18]. This hefty price tag is of major concern to the Arizona Department of Corrections which stated that the level of historical growth is unsustainable [18]. Many states are targeting transition and recidivism reduction initiatives as a more effective criminal justice investment than continuing to expand prison populations and construct new prison beds [18]. While it is clearly in the best interest of the Arizona Department of Corrections to reduce the rate of recidivism it is also in the best interest of the community;this would mean that the money saved could eventually be reinvested into other areas of the community that are underfunded such as education and infrastructure.

Arizona Department of Corrections published their 2018-2022 plan which includes a strategy to reduce recidivism through reentry preparation and support [19]. The current recidivism rate for Arizona inmates is 39.1% per year [20]. The strategies to reduce it include leading the state-level breakthrough project on recidivism reduction, improving inmate programs, increasing emphasis on inmate completions of degree, stabilizing inmates' mental health needs prior to release, increasing community engagement, and improving use of offender interventions and sanctions [20]. As an evaluation, the RAND (Research and Development) corporation conducted a study that evaluated the effectiveness of correctional education. Among the key findings, they found that correctional education improves inmates' chances of not returning to prison. Inmates who participate in correctional education programs had 43% lower odds of recidivating than those who did not, this translates to a reduction in the risk of recidivating of 13 percentage points. In other words, providing correctional education can be cost-effective when it comes to reducing recidivism [5].

In this study we focus on two types of educational programs; the first is the GED program offered in all of the state prisons while the second is the Reentry Court Program. The GED is offered to inmates who have completed the Functional Literacy program, an equivalence to having an 8th grade education [12]. The Reentry Court Program helps a parolee find a job and even offers temporary housing. In this study, we propose a mathematical approach to examine under what conditions the available program to inmates and the one available to parolees are the most cost effective and reduce recidivism in Arizona.

The goal of the mathematical study is to examine the qualitative differences between the GED and non-GED holders in prison so we can understand how to best invest monetary funds. Mechanically, inmates go through the same process but in order to really establish the effectiveness of education, we look at the qualitative journey which is different for each group. With an analogy to epidemiology where disease transmission is modeled, we use a compartmental and deterministic model to analyze mechanisms inside a given prison and inmates' influences on transmission of education as an infectious social disease where "education" is the disease. In this study, we only consider the Arizona population that is incarcerated in state prisons and exclude the private prisons because we don't have enough access to enough data. For simplicity we analyze the inmates with or without GED, with or without the transition program, and with or without recidivating.

A previous study in 2012 by Alvarez*et al.* [1] used a mathematical model to analyze how recidivism rates were impacted based on which reform programs the inmate attended. They concluded that recidivism rates lowered if inmates had the opportunity to go through both the inside and the outside programs. The study only analyzed data from California therefore has some geographical limitations. However, it did not employ a cost analysis of the economic impact that these programs had on the corrections budget.

In 2015, Purtolas *et al* modeled the direct impact of incentivized educational programs on the recidivism rates in Louisiana [6]. The study focused on finding out how much incentive the state has to provide in order to reduce recidivism. The claim is that in order for a prison's optimal profit strategy to reduce recidivism then an incentive has to be offered. The study found that more effective reform programs have a more cost effective strategy in reducing recidivism. However, the study found that the way the current prison system operates the prison is highly motivated to reduce the effectiveness of the reform programs.

Our model builds on the studies previously conducted and will use many of the same concepts in order to model recidivism in Arizona and conduct a cost analysis. Furthermore, our study places great emphasis on the available transition program that inmates have access once they are released.

We will consider programs that occur outside and inside prison. The program we focus on is the GED program that takes place inside prison where inmates can take classes in order to earn their degree and the transition program outside of prison that helps paroled inmates reintegrate into society. The GED program prepares inmates for life outside of prison and they are also a prerequisite for them to participate in the work programs offered inside of prison [12]. The GED program is mostly funded through the state except for the testing that comes at the inmates' expense. Both programs are completely optional to the parolee and they do not have to complete them. Studies have shown that inmates who complete both programs are less likely to return to prison [14]. The analysis of the model shows effectiveness of the education program inside and the transition program outside. We hypothesize that in order to reduce the levels of recidivism, the inmates of the Arizona state prisons would need to pass through the prison programs offered by the state during their sentence and after they are released. The following sections developed the ideas discussed here. In section 2, we discuss the methods used to develop our mathematical model and the dynamics of our model. In section 3, we focus on calculating the parameters for our model and estimating the ones for which we have limited data. Section 4 focuses on the analysis of the tools we developed that helped our study. We begin by discussing how we found the education free equilibrium and how we developed the cost function. Finally, section 5 explores the results we found by analyzing the education free equilibrium, the equilibrium points associated with the recidivist class, and the results of the cost equation. Section 6 will conclude our study and state our conclusions.

2 Methods

2.1 Model

The proposed model employs a system of ordinary differential equations, tracking the number of inmates within one prison that enter without GED and the number of prisoners with GED or higher educational degrees in a single prison. We then look at the likelihood of the inmate to recidivate once he is released.Our main objective is to find under what conditions are the available education programs, such as the GED and Reentry centers in Arizona, the most cost effective in reducing recidivism. We hypothesize that in order to reduce the levels of recidivism, the inmates of the Arizona state prisons would need to pass through the prison programs offered during and after their prison sentence.

2.2 Prison and Education Dynamics

In our model, the inside prison dynamics are given by the first-time offenders class which is divided into two parts: I_1 represents the first-time inmates that come in to prison without a GED and I_2 are those that are in prison for the first-time with a GED or higher education levels. The E compartments represent the inmates that have come back to prison after having already been freed once. These compartments are also divided into two parts: E_1 repeat offenders without GED, and E_2 repeat offenders with GED or more. These compartments allow us to keep the first time offenders separate from the ones that recidivized to help us calculate the cost. The transition rates between compartments on the inside are defined by contact rates between the uneducated and educated, assuming that the educated influence the uneducated to complete their GED. In the T compartments which are outside prison, T_1 represents the released inmates that were not involved in the GED program, but are involved in the reentry program. Class T_2 are the inmates that either have a GED or received one while inside and go to the reentry program. The O compartments are divided into: O_1, the released inmates not currently in a transitional program (and still at risk); and O_2, the released inmates that completed the GED program and are not currently in a transitional program (and still at risk). Even though the T and O compartments appear to be similar, the proportion of released inmates that exit from the T compartment are our ideal released inmate that leave the reentry program and never commit another crime. The O compartments are designed to catch the proportion of offenders that remain at risk of committing another crime. As a result, the exit rate from the T compartments is larger than the O compartments because we assume that the reentry program has a higher rate of effectiveness.

The proposed model assumes a constant recruitment rate (Λ_i, i=1,2) for each of the incoming populations of inmates (I_i), taking into account the incoming inmates with no GED, and the incoming inmates with a level of education of GED or more. It is assumed that a proportion (q_i) of inmates who leave prison after $\frac{1}{\alpha_i}$ days will go through the transition program (T_i), while (1 -q_i) remain at risk of recidivism (O_i). Among those released inmates, a fraction $\left(\frac{\mu_{1,2}}{\gamma_i+\mu_{1,2}}\right)$ will be rehabilitated due to the effectiveness of the program, while the rest $\left(\frac{\gamma_i}{\gamma_i+\mu_{1,2}}\right)$ will remain at risk of recidivism. While the inmates are at risk of recidivism, they can go back to prison at a per capita rate p_i or they can rehabilitate at per capita rate $\mu_{3,4}$.

Our model looks at how peer influence between both of these groups (the group with no GED and the group with GED or more) affect the population density in the different compartments. The incoming population of inmates with no GED is represented by Λ_1 and the incoming population of inmates with GED or more is represented by Λ_2. Even though both of the prisoner classes can attend a transition program when released, released inmates with high education levels have less chance of recidivism [5]. In this case, we are assuming that the infection represents the positive influence that the population with GED or more has over the population with no GED.

The classes and parameters of the model are explained in further detail in Tables 1, 2.

Table 1: Definition of variables in the model

Class	Description
I_1	Inmates with no GED
I_2	Inmates with GED or more
E_1	Returning offenders without GED
E_2	Returning offenders with GED or more
T_1	Released inmates without GED attending the transition program
T_2	Released inmates with GED attending the transition program
O_1	Inmates without GED who are still at risk to recidivate
O_2	Inmates with GED who are still at risk to recidivate

Table 2: Description of the parameters used in the model

Parameters	Description	Units
Λ_1	Incoming population without GED	People per time
Λ_2	Incoming population with GED or more	People per time
β_1	Rate of influence between I_2 and I_1	Per person per time
β_2	Rate of influence between E_2 and E_1	Per person per time
p_1	Rate of people leaving O_1	1/time
p_2	Rate of people leaving O_2	1/time
q_1	Proportion of released inmates without GED who go to the transition program	Dimensionless
q_2	Proportion of released inmates with GED or more who go to the transition program	Dimensionless
μ_1	Per capita rehabilitation rate from T_1	1/time
μ_2	Per capita rehabilitation rate from T_2	1/time
μ_3	Per capita rehabilitation rate from O_1	1/time
μ_4	Per capita rehabilitation rate from O_2	1/time
γ_1	Rate of people leaving T_1	1/time
γ_2	Rate of people leaving T_2	1/time
α_1	Rate of people leaving I_1	1/time
α_2	Rate of people leaving I_2	1/time

The system of ordinary differential equations (1) captures the aforementioned prison education dynamics. It is also represented in Figure 1

$$
\begin{aligned}
\frac{dI_1}{dt} &= \Lambda_1 - \beta_1 I_1(I_2+E_2) - \alpha_1 I_1 \\
\frac{dI_2}{dt} &= \Lambda_2 + \beta_1 I_1(I_2+E_2) - \alpha_2 I_2 \\
\frac{dE_1}{dt} &= p_1 O_1 - \beta_2 E_1(I_2+E_2) - \alpha_1 E_1 \\
\frac{dE_2}{dt} &= p_2 O_2 + \beta_2 E_1(I_2+E_2) - \alpha_2 E_2 \\
\frac{dT_1}{dt} &= (I_1+E_1)(q_1\alpha_1) - T_1(\mu_1+\gamma_1) \\
\frac{dT_2}{dt} &= (I_2+E_2)(q_2\alpha_2) - T_2(\mu_2+\gamma_2) \\
\frac{dO_1}{dt} &= \gamma_1 T_1 + (1-q_1)(\alpha_1 I_1 + \alpha_1 E_1) - O_1(p_1+\mu_3) \\
\frac{dO_2}{dt} &= \gamma_2 T_2 + (1-q_2)(\alpha_2 I_2 + \alpha_2 E_2) - O_2(p_2+\mu_4)
\end{aligned} \tag{1}
$$

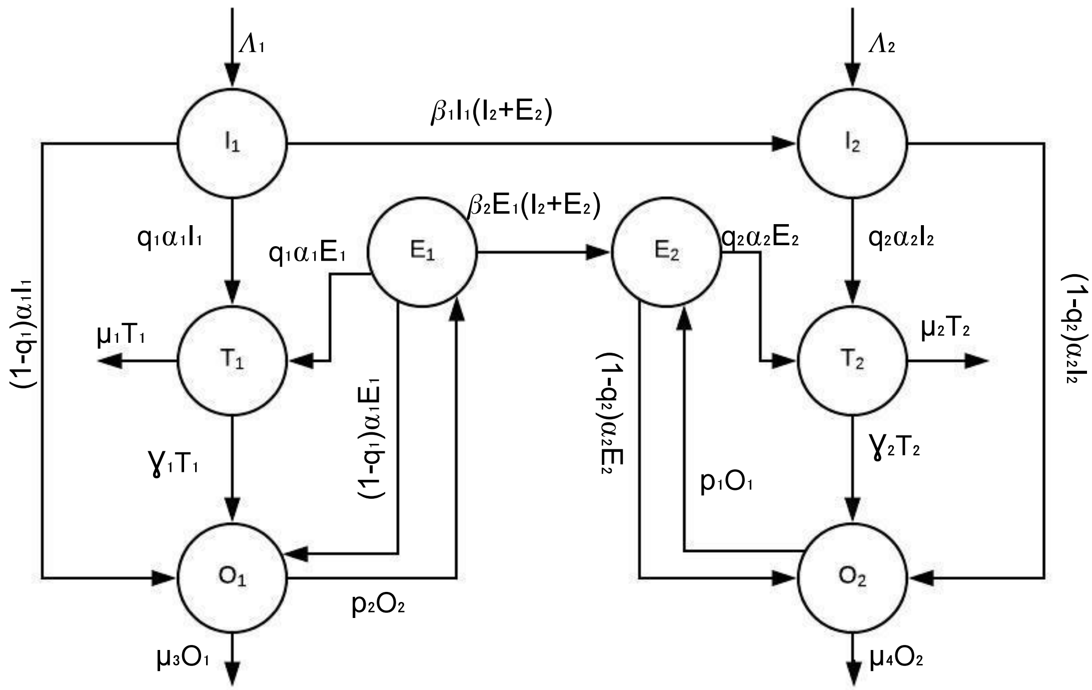

Figure 1: Model

3 Parameters

Our parameters as defined in Table 2 can mostly be calculated from the data that have been gathered during our investigation. When calculating the parameters, we only focused on one prison therefore the total population comes from the mean of the total population of the Florence State Prison which is $3,955$ [11]. Our rates of movement from one compartment to another are rescaled to fit this population. We also focused on the male population older than 18 years of age. While some values were calculated in a pretty straightforward manner others required estimation as shown below.

3.1 Parameter Calculations

On calculating α_1 and α_2, we supposed that the exit rates are the same regardless of GED. As stated on the Arizona Department of Corrections website, the average prison duration term is 24 months [11]. Therefore:

$$\alpha_1 = \alpha_2 = \frac{1}{\text{Average Length of Stay}} = \frac{1}{24} = 0.04 \text{ /month}$$

The direct data for Arizona results were not found for the number of inmates coming in with a GED. Therefore we used the 2016 data from Florida's Department of Corrections statistics on TABE, the test of adult basic education, to calculate Λ_i, the incoming population with a GED [15]. We decided to use data from Florida since they have data available about their prisoner's education level. We find that 44% of inmates that are being admitted, have no priors, and have GED Literacy [16]. We then calculated the rate at which male inmates with no priors are being admitted in Arizona. This will be an estimated calculation since we used data from Florida, since .

$$\frac{\text{Males without priors admitted}}{year} = \frac{x}{\text{month}},$$

$$\frac{\text{10,236 people}}{\text{12 months}} = \frac{x}{\text{1 month}},$$

$$x = 853 \text{ admissions per month},$$

$$\frac{\text{853 people}}{\text{1 month}} \times .44 = \frac{\text{478 people}}{\text{1 month}},$$

$$\Lambda_2 = 478 \text{ people per month},$$

$$\Lambda_1 = 375 \text{ people per month}.$$

p_1 represents the rate of recidivism of those without GED. We calculate this rate from a recidivism study done in Arizona in 2005 [10] which shows that 42.4% of parolees returned to the Arizona Department of Corrections custody for any reason over a 3 year period. Using this we calculate the rate of recidivism in Arizona. We will rescale this percentage to reflect our month scale.

$$p_1 = \frac{\text{Recidivism Rate}}{\text{3 years}},$$

$$p_1 = \frac{.424}{\text{36 months}},$$

$$p_1 = 0.0118 \text{ per month}.$$

In order to calculate p_2 (the rate of recidivism of those with GED or more) we consider the same study used to derive p_1. The study found that those who participate in inmate programs like academic education reduce the recidivism rates by an average of 25% [10]. As a result, we used the same calculations but using a recidivism rate of 31.8% which is 25% lower.

$$\text{New percentage} = \text{recidivism rate} - (\text{recidivism rate} \times \text{percentage lowered}) = \text{new percentage}$$

$$(.424 - (.424 \times .25)) \times 100 = 31.8\%$$

Now running through the same calculations as p_1 using our new percentage we find that

$$p_2 = \frac{.318}{\text{36 months}},$$

$$p_2 = 0.009 \text{ per month}.$$

3.2 Parameter Estimation

In this section we estimated the parameters β_i, μ_i, γ_i, and q_i since the meaning of these parameters involved various methods that allow them to fit our assumptions.

We will begin with the estimation of μ_1 and γ_1.The basic structure of the model creates two unique exit rates that represent the same thing; μ_1 is the exit rate from the transition program that assumes that the parolee will cease to recidivate, and γ_1 is incorporating the latency period where the parolee can still recidivate and become an inmate again. As a result, the calculated rate at which a parolee leaves the transition program will be split amongst the unknown probability of the program success. We begin by calculating the rate at which people leave the transition program which we will call η.

As mentioned before the main difference between our study and previously conducted studies is that we analyzed the impact of having a transition program once the inmates are released. Currently, Arizona has the Maricopa Reentry Center and the Pima Reentry center. According to a table presented in February 2018 to the Appropriations Committee by the Arizona Department of Correction we calculate η, the rate at which people leave the transition program [13]. The transition program is currently, at a maximum, 90 days long.

$$\eta_1 = \frac{1}{\text{3 months}},$$

$$\eta_1 = 0.3 \text{ per month},$$

Now in order to incorporate this exit rate properly, we introduced the variable a which will be the probability of the program's success, a number for which we have no data. Using this a we can define μ_1 and γ_1.

$$\mu_1 = a * \eta,$$

$$\gamma_1 = (1 - a) * \eta.$$

We multiply a by μ_1 with the assumption that the program will have a slightly greater success. In order to begin the process, we will arbitrarily let $a = \frac{1}{e}$ which is about a 37% success rate. Therefore,

$$\mu_1 = 0.12 \text{ per month},$$

$$\gamma_1 = 0.21 \text{ per month}.$$

In this case, we will suppose that $\gamma_1 = \gamma_2$. Therefore $\gamma_2 = 0.21$ per month.

The rest of the exit rates μ_2, μ_3, and μ_4 are defined with respect to μ_1.

Assumption 1: GED parolee's have a double chance of reforming completely. This yields the following result:

$$\mu_2 = 2\mu_1 = 0.24 \text{ per month}$$

Assumption 2: Non-GED parolees have half the chance of reforming without the transition program.

$$\mu_3 = \tfrac{1}{2}\mu_1 = 0.06 \text{ per month}$$

Assumption 3: GED parolees that do not go through the GED program have the same chance as non-GED parolees that do go through the transition program.

$$\mu_4 = \mu_1 = 0.12 \text{ per month}$$

In order to calculate β_1, we gathered the data of total monthly enrollments for the GED program from the Arizona Department of Corrections reports [11]. We compiled the data into a list of data points and created a graph of the overall enrollment for a total of three years which is typically the length of time used to measure recidivism.

In our model we assume that the social influence, denoted as β_1, is the only way other inmates are motivated to sign up for the GED classes. Therefore, we expect β_1 to reflect our collected data. In order to achieve this, we used a Monte Carlo Fitting on our data. We consider a simple S-I-R model to get an approximate value of β_1 that will best estimate the data that we have currently. We also supposed that the length of time that the infected individual has to influence the susceptible one is the average release rate [11]. After running our simulation of the SIR model under mass incidence conditions, we found that $\mathcal{R}_e = 1.175$ (see fig. 2). In our simplified model $\mathcal{R}_e = \frac{N\beta}{\gamma}$. In our model $\gamma =$ average rate of release for inmate $= 0.04$ per month, $N =$ Total Population of prison. Therefore $\beta = \frac{\gamma * R_e}{N} = 0.000012$ per month.

In our model, the β effect is split between the I classes and the E classes but the β that was calculated using the simplified version was not. As a result, we make the assumption that the recidivists will have less motivational powers on other inmates which can be expressed as $\beta_1 > \beta_2$. Since we have no available data for social influence, the division of our calculated β will be completely arbitrary.

$\beta_1 = \beta - \beta_2 =$
0.000007 per month,

$\beta_2 = \beta - \beta_1 =$
0.000005 per month,

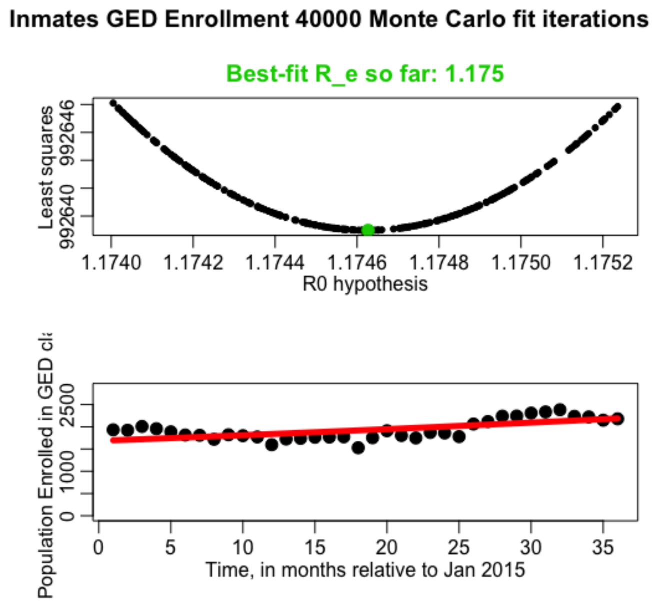

Figure 2: β estimation

Finally, we estimated the values of q_1 and q_2, the proportion of released GED and non-GED holders enrolling in a reentry program. In this case we look at the Maricopa Transition Center which has a maximum capacity of 100 beds and maximum length of stay of 90 days. We begin by calculating the number of beds available per month.

$$\text{Beds Available per month} = \frac{\text{100 beds}}{\text{3 months}} = \frac{\text{33 beds}}{\text{1 month}}$$

Then we looked at the average number of people released per month and multiplied it by the number of beds available per month to get our maximum proportion of q.

$$q_m = \frac{\text{33 beds}}{\text{1 month}} \times \frac{1}{\text{126people}} = 0.3$$

In order to calculate q_1 and q_2 we define a new equation that takes into account the flow rates of inmates. Therefore if there are very few people from I_1 and. E_1 released then more inmates from I_2 and E_2 will be recruited.

$$q_1 \times \alpha_1 + q_2 \times \alpha_2 = q_m \times (\alpha_1 + \alpha_2)$$

We make the assumption that inmates without a GED will need more help and therefore we pick the numerical value of q_2 and let that define q_1 which will be larger than q_m

$$q_2 = 0.3$$

$$q_1 = 0.4$$

4 Analysis

4.1 R_e Analysis

After conducting model analysis under our starting assumptions, we found that our model has an education-free equilibrium (EFE) (see appendix A). In order to measure the strength of peer influence on the education rate, we consider the special case where $\Lambda_2 = 0$, i.e., no one comes in educated. Only then we find our EFE (in which $I_2^* = E_2^* = T_2^* = O_2^* = 0$) which yields the R_e, the basic reproductive number. Then we find the EFE values, including I_1^* and E_1^*, which appear in the expression 2 for R_e. The education reproductive number, R_e, is computed using the next generation operator method [22], with the following vectors:

$$\mathcal{F} = \begin{pmatrix} \beta_1 I_1 \left(E_2 + I_2\right) + \Lambda_2 \\ \beta_2 E_1 \left(E_2 + I_2\right) \\ 0 \\ 0 \end{pmatrix},$$

$$\mathcal{V} = \begin{pmatrix} \alpha_2 I_2 \left(-\left(1 - q_2\right)\right) - q_2 \alpha_{2 I_2} \\ -\alpha_2 E_2 \left(1 - q_2\right) - \alpha_2 E_2 q_2 + O_2 p_2 \\ \alpha_2 E_2 q_2 + \alpha_2 I_2 q_2 + \left(-\gamma_2\right) T_2 - \mu_2 T_2 \\ \alpha_2 E_2 \left(1 - q_2\right) + \alpha_2 I_2 \left(1 - q_2\right) - \mu_4 O_2 - O_2 P_2 + \gamma_2 T_2 \end{pmatrix}$$

Here, the $\mathcal{F}$ vector is composed of the rates at which new inmates are admitted into our initial compartments, and the $\mathcal{V}$ vector has the transfers of individuals between compartments. We are then able to compute the respective matrices F and V:

$$F = \begin{pmatrix} \beta_1 I_1 & \beta_1 I_1 & 0 & 0 \\ \beta_2 E_1 & \beta_2 E_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, V = \begin{pmatrix} \alpha_2 & 0 & 0 & 0 \\ 0 & \alpha_2 & 0 & -p_2 \\ -\alpha_2 q_2 & -\alpha_2 q_2 & \gamma_2 + \mu_2 & 0 \\ \alpha_2 \left(q_2 - 1\right) & \alpha_2 \left(q_2 - 1\right) & -\gamma_2 & \mu_4 + p_2 \end{pmatrix}$$

Following the last steps of the next generation matrix [22](all the steps are shown in Appendix A.1), we found a matrix with 2 rows that are multiples of each other, and the 2 others are 0 rows:

$$\begin{pmatrix} \frac{I_1\beta_1}{\alpha_2} + \frac{I_1 p_2(\gamma_2 - q_2\mu_2 + \mu_2)\beta_1}{\alpha_2(p_2 q_2 \mu_2 + (\gamma_2+\mu_2)\mu_4)} & \frac{I_1\beta_1(\gamma_2+\mu_2)(p_2+\mu_4)}{\alpha_2(p_2 q_2\mu_2+(\gamma_2+\mu_2)\mu_4)} & \frac{I_1 p_2 \beta_1 \gamma_2}{p_2 q_2 \alpha_2 \mu_2 + \alpha_2 \mu_4 \mu_2 + \alpha_2 \gamma_2 \mu_4} & \frac{I_1 p_2 \beta_1(\gamma_2+\mu_2)}{\alpha_2(p_2 q_2\mu_2+(\gamma_2+\mu_2)\mu_4)} \\ \frac{E_1\beta_2}{\alpha_2} + \frac{E_1 p_2(\gamma_2 - q_2\mu_2 + \mu_2)\beta_2}{\alpha_2(p_2 q_2 \mu_2 + (\gamma_2+\mu_2)\mu_4)} & \frac{e_1\beta_2(\gamma_2+\mu_2)(p_2+\mu_4)}{\alpha_2(p_2 q_2\mu_2+(\gamma_2+\mu_2)\mu_4)} & \frac{E_1 p_2 \beta_2 \gamma_2}{p_2 q_2 \alpha_2 \mu_2 + \alpha_2 \mu_4 \mu_2 + \alpha_2 \gamma_2 \mu_4} & \frac{E_1 p_2 \beta_2(\gamma_2+\mu_2)}{\alpha_2(p_2 q_2\mu_2+(\gamma_2+\mu_2)\mu_4)} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The education reproduction number is therefore given by the largest eigenvalue of the matrix above. After some simplifications we find:

$$R_e = \frac{I_1^* \beta_1 + E_1^* \beta_2}{\alpha_2 (1 - \frac{p_2}{p_2+\mu_4}(1 - q_2 * \frac{\mu_2}{\gamma_2+\mu_2}))} \tag{2}$$

$$R_e = \frac{I_1^*\beta_1 + E_1^*\beta_2}{\alpha_2} \left(\frac{1}{1 - \frac{p_2}{p_2+\mu_4}[1 - q_2(\frac{\mu_2}{\gamma_2+\mu_2})]} \right)$$

All the calculations are shown in Appendix A.1

In order to explain what we got, we define $r_2 = [\frac{p_2}{p_2+\mu_4}(1 - q_2 * \frac{\mu_2}{\gamma_2+\mu_2})]$ which can be interprent as the proportion of released inmates who recidivate. The term $\frac{p_2}{p_2+\mu_4}$ is the proportion of people from O_2 who go to E_2, multipliying by $(1 - q_2 * \frac{\mu_2}{\gamma_2+\mu_2})$ which is the proportion of people who go to O_2. This path can become a loop, and furthermore if we take into account this whole terms that make up the denominator, we will realize that it behaves as a geometric series. The behavior of the terms depend on the ratio r_2, which can be understood as the proportion of released inmates who recidivate . Which turns out to give us all the different paths from E_2 that an inmate can take in order to go back to prison, then:

$$R_e = \frac{I_1^*\beta_1 + E_1^*\beta_2}{\alpha_2}\left(\frac{1}{1-r_2}\right) \tag{3}$$

At the end, the R_e has the term that includes r_2, which represents the recidivate proportion being multiplying by the influence of inmates with GED to the other who don't have it, divided by the frequency or rate that this happens. In addition, to analyze the stability conditions for R_e, we used the criteria for R_e [7]. We need to keep our reproductive number $R_e > 1$ in order to keep spreading the education transmission, using the non-simplified form of our R_e because is more simple. Also, it can be found explicitly in Appendix A.1 :

$$R_e = \frac{(\beta_1 I_1^* + \beta_2 E_2^*)(\gamma_2 + \mu_2)(p_2 + \mu_4)}{\alpha_2(p_2 q_2 \mu_2 + \mu_4(\gamma_2 + \mu_2))}$$

The basic reproductive number R_e of our discrete model is defined and the dynamical behavior of the model is studied. It is proved that the education free equilibrium, which is only possible if $\Lambda_2 = 0$ and is globally asymptotically unstable if $R_e < 1$, and the persistence of the model is obtained when $R_e > 1$. It is necessary to focus in the global stability of the endemic equilibrium. Sufficient conditions for the global stability of the endemic equilibrium are established by using the comparison principle. Numerical simulations are done to show our theoretical results and to demonstrate the complicated dynamics of the model1. R_e is the average number of secondary infections produced by one infected individual during the entire course of infection in a completely susceptible population, in our case the infection is the education. R_e often serves as a threshold parameter that predicts whether an infection dies out or keeps persistence in a population [9].
Therefore, for our model the persistence of education influence is determined by the stability of the disease free equilibrium and the existence of endemic equilibrium of model1. R_e is always positive if:

$$(\beta_1 I_1^* + \beta_2 E_2^*) < \frac{\alpha_2(p_2 q_2 \mu_2 + (\gamma_2 + \mu_2)\mu_4)}{(\gamma_2 + \mu_2)(p_2 + \mu_4)}$$

By forthright and mindful calculations shown in Appendix A.2, we know that model 1 has a unique endemic equilibrium when $\Lambda_2 > 0$ and $R_e > 1$.

4.2 Cost function

The total cost incurred in implementing control measures by the education programs in the Arizona Department of corrections is modeled by the function

$$C = c_1(E_1^* + E_2^*) + c_2(T_1^* + T_2^*) + c_3(E_2^* + I_2^*) \tag{4}$$

Where c_1 denotes the average cost incurred by the state of Arizona to maintain an incarcerated person per day, c_2 represents the cost incurred by the state of Arizona to maintain a former prisoner in a non-residential transition program and c_3 is the amount that is spent by the state of Arizona in GED classes per prisoner daily [8]. According to the Arizona Department of Corrections Per Capita Cost Report of FY 2017, it costs the state \$ 68.55 per day to maintain an inmate. While inside the education classes available for inmates cost the state an average of \$ 2.25 per day. The cost to maintain per inmate in a non-residential transition center is \$ 10.71 per day [4].

Since the parameters are measured in months, we need to convert this values of per person per day into per person per month:

$$\begin{aligned}
c_1 &= \frac{\text{cost/day} * 365}{12} &&= \frac{68.55 * 365}{12} = 2085.06 \text{ per person per month}\\
c_2 &= \frac{\text{cost/day} * 365}{12} &&= \frac{10.71 * 365}{12} = 325.76 \text{ per person per month}\\
c_3 &= \frac{\text{cost/day} * 365}{12} &&= \frac{2.25 * 365}{12} = 68.44 \text{ per person per month}
\end{aligned}$$

To estimate the numbers for the values of $E^*_1,_2$, $T^*_1,_2$, and I^*_2, we would need to use the endemic equilibrium equations and their numerical values, this section can be found on Appendix A.2. When we evaluate for each parameter we find that the value of E_1 is ≈ 89 $E_2 \approx 704$, $T_1 \approx 98$, $T_2 \approx 372$, and $I_2 \approx 19{,}237$. Replacing these values in the original equation 4 we get:

$$C = 2085.06(89 + 704) + 325.76(98 + 372) + 68.44(704 + 19237) \approx \$3,171,900 \tag{5}$$

We want to consider how the participation rates in transition programs affect this cost, so we write our function C as:

$$C = f(q_1, q_2) \tag{6}$$

The concept of cost-effectiveness is used to compare strategies in terms of cost per inmate education achieved in implementing a particular strategy. In this case an inmate has the opportunity to get education classes while in prison and go through a transition program. The aim is to find which strategy is the most cost-effective when the Department of Corrections of Arizona is willing to spend a certain amount per unit increase in effectiveness. In general, finding the most cost-effective model is a two-step process [8]. First, we need to check if offering inside education only is most cost effective compared to offering inside education and the outside program and second, we try to find the most cost-effective strategy among all cost-effective strategies. Although the inside education program only strategy may prove more cost effective, it may not be as effective as our second option over a certain period of time since data suggest higher levels of recidivism for people who have only taken only education programs inside [1]. The inmates who take the education programs inside and when released go through the transition program seem to have a large cost. However, further studies have suggested that it is efficient in lowering recidivism [1].

Only counting the inside program we find from 2017 data:

$$C = c_1(E^*_1 + E^*_2) + c_3(E^*_2 + I^*_2) = 2085.06(89 + 704) + 68.44(704 + 19237) \approx \$3,018,210 \tag{7}$$

5 Results

In this section we provide the discussion, due to the selection process used we attempted to achieve a cost analysis for the implementation of Education programs with 2017 data. In addition we also obtained a numerical value for the reproductive number and we show numerical simulations of the model.

5.1 Numerical Analysis of R_e

Using the parameter calculations and the values for the variables found in the appendix we assigned numerical values to I_1,E_1,β_1, β_2, γ_2, μ_2, μ_4, α_2. This allowed us to create a function for R_e in order to analyze the effects of our parameters p_2, the recidivism rate of educated prisoners, and q_2, the proportion of educated inmates that enter the GED program. The function for R_e is as follows

$$R_e = \frac{0.310375}{1 - \frac{p_2 * (1 - 0.827586 * q_2)}{0.12 + p_2}} \tag{8}$$

We then graphed a contour plot and found the effect of recidivism on education. We found that as the recidivism rate increases R_e becomes greater than 1. The proportion of inmates that have a GED and leave the prison has to be below 0.35 otherwise R_e falls below 1. In our model, we want R_e to be greater than 1 because our infection is education and the goal is for people to get educated.

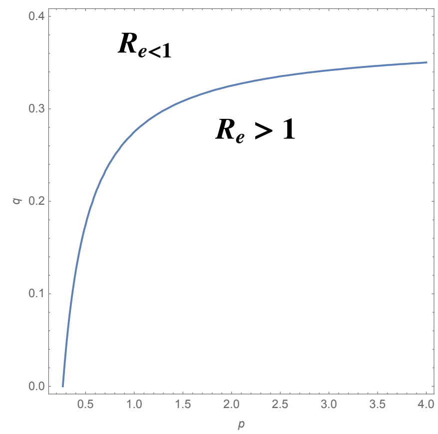

Figure 3: R_e Contour Plot

5.2 Equilibrium Analysis

We now look at the effect of education on recidivism. In order to observe this, we look at the two compartments that compose the recidivist class.

$$R = E_1 + E_2$$

In order to analyze the recidivist class we used the equilibrium points:

$$E_1^* = \frac{\mathrm{r1}(\alpha 2(-\mathrm{i2})+\lambda 1+\lambda 2)}{\frac{\alpha 1\beta 2(\alpha 2\mathrm{i2}-\lambda 2)}{\beta 1(\alpha 2(-\mathrm{i2})+\lambda 1+\lambda 2)}+\alpha 1(1-\mathrm{r1})}$$

$$E_2^* = \frac{\alpha 1(\alpha 2\mathrm{i2}-\lambda 2)}{\beta 1(\alpha 2(-\mathrm{i2})+\lambda 1+\lambda 2)} - I_2^*$$

I_2^* is located in the Appendix A.2.

The following graph of this function gives us the behavior of the recidivism independent of the initial conditions of our system.

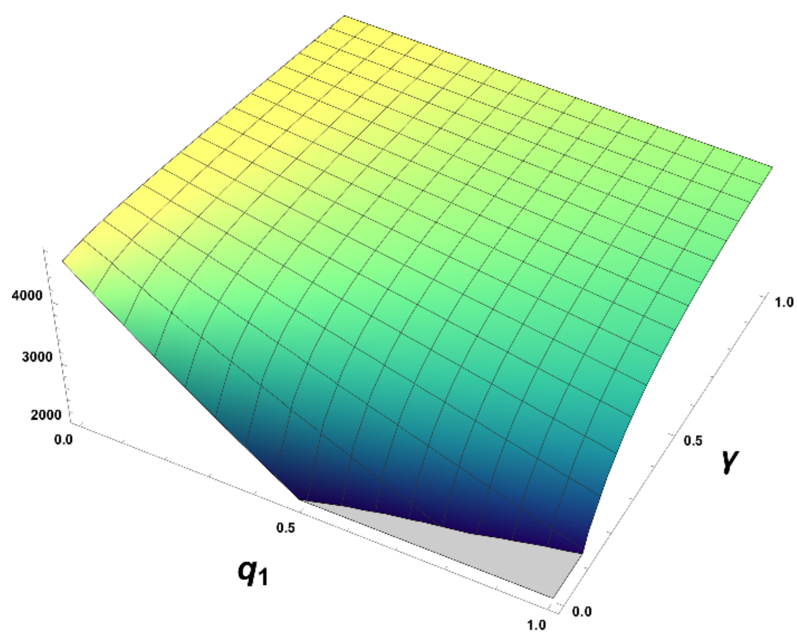

Figure 4: R_e Recidivism Behavior

We begin by looking at the behavior of γ which represents the length of the transition program. Since we have an inverse proportion, we see that as the length of the transition program increases, the recidivist population is reduced.

For the proportion of uneducated inmates, we see that as the proportion of uneducated inmates going into the transition program gets larger, the recidivist population is also reduced.

Our graph also shows a point where the proportion of uneducated inmates leaving the prison and entering the transition program combined with a long enough program will make it possible to completely remove recidivism.

5.3 Cost Analysis

The cost function is $C = c_1(E_1^* + E_2^*) + c_2(T_1^* + T_2^*) + c_3(E_2^* + I_2^*)$. This represents the sum of the cost of recidivism ($c_1(E_1^* + E_2^*)$, cost of training program outside the jail($c_2(T_1^* + T_2^*)$), and the cost of GED inside the prison($c_3(E_2^* + I_2^*)$). This function depends on the number of released inmates(q_1 and q_2) going through the transition program.
If $q_1 = q_2 = 1$ (all released prisoners go through the transition programs) , The total cost is:
$C = c_1(E_1^*+E_2^*)+c_2(T_1^*+T_2^*)+c_3(E_2^*+I_2^*) = 2085.06(66+416)+325.76(264+1745)+68.44(416+19211) =$ $\$3.00272 * 10^6$ Comparing to the original cost when q_1 and q_2 are different, there is \$169,180 saving.
If $q_1 = q_2 = 0$ (Absence of transition programs so $T_1 = T_2 = 0$), the total coast is:
$= 2085.06(104 + 808) + 68.44(808 + 19246) = \$3.27407 * 10^6$ We can notice that the cost is even more than the original, there is an increment of \$102,170.

Since the cost depends on the quantities q_1 and q_2, and our ultimate goal is to minimize the cost, we graphed the cost function while varying both quantities (q_1 and q_2). The following graph depicts this variation and shows where the cost is at its minimum.

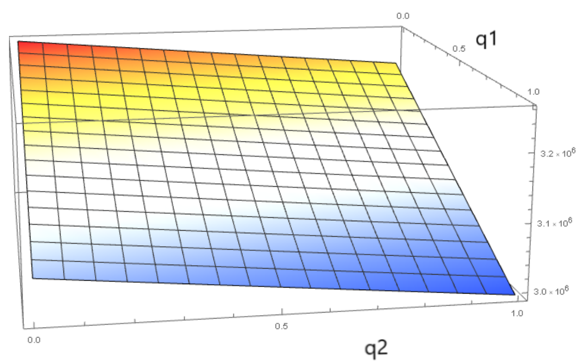

Figure 5: Total cost as function of the participation on re-entry programs

In figure 5, we can see that that cost function is at its minimum when $q_2 = q_1 = 1$. This can be translated to universal transition program where all released inmates go through it. According to our model, this will lead to a minimum cost comparing to other situations. In addition to that, this has the potential to generate more income since people with higher education earn more on average which will result to higher contribution to taxes and less dependency on government assistance [2].

6 Conclusion

Education is typically defined as something that occurs during youth it does not become a choice until you become an adult. Education doesn't need to be traditional it can also take on many different forms that involve learning. In our study we look at both aspects of education, the academic GED side which we suppose will give the inmate an advantage when it comes to job hunting, and the non-academic side or the transition program which creates opportunities for the parolee that will help the inmate avoid going to back to prison. Because of this, we have to look at R_e, the transmission of education, from a unique perspective. We assume that education would help a paroled inmate succeed and avoid returning to prison and with enough peer influence they will be convinced to join the educated inmates. We calculated R_e as being 1.175, which means that every educated inmate can influence at least 1 uneducated inmate. From the equilibrium points we reached the conclusion that it is possible to reduce recidivism to 0 given that enough uneducated inmates enter the transition program and the length of the program is long enough. Therefore we conclude that allocating more resources to the transition program will aid in helping the state of Arizona reduce recidivism.

In order to better make a statement about the cost we conducted a cost-analysis and found that transition programs in Arizona eventually pay for themselves because the total cost of is minimized when every released inmate enters a reentry program. This conclusion agrees with our conclusion above that reducing recidivism requires a large proportion of prisoners to enter the transition programs upon release. This proves our hypothesis that was postulated. Therefore we make the recommendation that more resources and funding needs to be allocated to transition programs. We also recommend that the state of Arizona conduct further studies in order to identify characteristics of successful transition programs in order to further strengthen the qualitative analysis of the programs.

There are other factors that affect recidivism like mandatory minimums, work-program agreements, and felony laws concerning sentencing. There are also demographic factors, length of time served, and type of offense that can be further explored within the population. Therefore, in order to strengthen our study we would want to incorporate more variables and parameters that can affect recidivism, not just education, and create a more exhaustive and comprehensive model that can truly help us understand the more complex dynamics of prisons.

7 Acknowledgments

We would like to thank Dr. Carlos Castillo-Chavez, Founding and Co-Director of the Mathematical and Theoretical Biology Institute (MTBI), for giving us the opportunity to participate in this research program. We would also like to thank Co-Director Dr. Anuj Mubayi as well as Coordinator Ms. Rebecca Perlin and Management Intern Ms. Sabrina Avila for their efforts in planning and executing the day to day activities of MTBI. We also want to give special thanks to Dr. Christopher M Kribs, Bechir Amdoundi, Dr. Naala Brewer, and Dr. Baltazar Espinoza. This research was conducted as part of 2018 MTBI at the Simon A. Levin Mathematical, Computational and Modeling Sciences Center (MCMSC) at Arizona State University (ASU). This project has been partially supported by grants from the National Science Foundation (NSF Grant MPS-DMS-1263374 and NSF Grant DMS-1757968), the National Security Agency (NSA Grant H98230-J8-1-0005), the Alfred P. Sloan Foundation, the Office of the President of ASU, and the Office of the Provost of ASU.

A Appendix

A.1 Appendix A: Education free equilibrium and R_e

This equilibrium only exits if we are assume $\Lambda_2 = 0$. Then, considering all the 8 equilibrium conditions, we got as a result that $E_2, O_2, T_2, I_2 = 0$ in order to stop the transmission: So, the remaining equations are:

1. $\frac{dI_1}{dt} = \Lambda_1 - \alpha_1 I_1$

2. $\frac{dE_1}{dt} = p_2 O_1 - \alpha_1 E_1$

3. $\frac{dT_1}{dt} = q_1 \alpha_1 (I_1 + E_1) - T_1(\mu_1 + \gamma_1)$

4. $\frac{dO_1}{dt} = \gamma_1 T_1 + (1 - q_1)\alpha_1 (E_1 + I_1) - O_1(\mu_3 + p_1)$

Using Mathematica we set the equations to zero and solve for the Education free equilibrium(EFE). After this, we get the following values and we added at star as an exponent to differentiate them.

$$\begin{aligned}
I_1^* &= \frac{\Lambda_1}{\alpha_1} \\
E_1^* &= \frac{\Lambda_1}{\alpha_1} p_1 \frac{\mu_1(1 - q_1) + \gamma_1}{\mu_3(\gamma_1 + \mu_1) + \mu_1 q_1 p_1} \\
T_1^* &= \Lambda_1 \frac{q_1(p_1 + \mu_3)}{\mu_3(\mu_1 + \gamma_1) + p_1 q_1 \mu_1} \\
O_1^* &= \Lambda_1 \frac{\mu_1(1 - q_1\mu_1) + \gamma_1}{\mu_3(\gamma_1 + \mu_1) + p_1 q_1 \mu_1}
\end{aligned}$$

To compute the basic reproductive number, we use the next generation operator. First, from the original equations of our model we create the $\mathcal{F}$ vector that is based on new criminals infections that will help to find the F matrix

$$\mathcal{F} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} I_1\beta_1 (I_2 + E_2) + \Lambda_2 \\ E_1\beta_2 (I_2 + E_2) \\ 0 \\ 0 \end{pmatrix}$$

Then, we compute the F matrix

$$F = \begin{pmatrix} \frac{\partial(x_1)}{\partial(I_2)} & \frac{\partial(x_1)}{\partial(E_2)} & \frac{\partial(x_1)}{\partial(T_2)} & \frac{\partial(x_1)}{\partial(O_2)} \\ \frac{\partial(x_2)}{\partial(I_2)} & \frac{\partial(x_2)}{\partial(E_2)} & \frac{\partial(x_2)}{\partial(T_2)} & \frac{\partial(x_2)}{\partial(O_2)} \\ \frac{\partial(x_3)}{\partial(I_2)} & \frac{\partial(x_3)}{\partial(E_2)} & \frac{\partial(x_3)}{\partial(T_2)} & \frac{\partial(x_3)}{\partial(O_2)} \\ \frac{\partial(x_4)}{\partial(I_2)} & \frac{\partial(x_4)}{\partial(E_2)} & \frac{\partial(x_4)}{\partial(T_2)} & \frac{\partial(x_4)}{\partial(O_2)} \end{pmatrix} = \begin{pmatrix} \beta_1 i_1 & \beta_1 i_1 & 0 & 0 \\ \beta_2 e_1 & \beta_2 e_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

From the terms of the original equations that are left, we are able to create the $\mathcal{V}$ vector, which follows the inflow and outflow of criminals from each compartment

$$\mathcal{V} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} -q_2\alpha_{2I_2} - (1-q_2)\,\alpha_2 I_2 \\ p_2O_2 - (1-q_2)\,\alpha_2E_2 - q_2\alpha_2E_2 \\ -\gamma_2T_2 - \mu_2T_2 + q_2\alpha_2I_2 + q_2\alpha_2E_2 \\ (1-q_2)\,\alpha_2I_2 - \mu_4O_2 + \gamma_2T_2 - P_2O_2 + (1-q_2)\,\alpha_2E_2 \end{pmatrix}$$

Here is the computation of the V matrix

$$V = \begin{pmatrix} \frac{\partial(y_1)}{\partial(I_2)} & \frac{\partial(y_1)}{\partial(E_2)} & \frac{\partial(y_1)}{\partial(T_2)} & \frac{\partial(y_1)}{\partial(O_2)} \\ \frac{\partial(y_2)}{\partial(I_2)} & \frac{\partial(y_2)}{\partial(E_2)} & \frac{\partial(y_2)}{\partial(T_2)} & \frac{\partial(y_2)}{\partial(O_2)} \\ \frac{\partial(y_3)}{\partial(I_2)} & \frac{\partial(y_3)}{\partial(E_2)} & \frac{\partial(y_3)}{\partial(T_2)} & \frac{\partial(y_3)}{\partial(O_2)} \\ \frac{\partial(y_4)}{\partial(I_2)} & \frac{\partial(y_4)}{\partial(E_2)} & \frac{\partial(y_4)}{\partial(T_2)} & \frac{\partial(y_4)}{\partial(O_2)} \end{pmatrix} = \begin{pmatrix} \alpha_2 & 0 & 0 & 0 \\ 0 & \alpha_2 & 0 & -p_2 \\ -\alpha_2q_2 & -\alpha_2q_2 & \gamma_2+\mu_2 & 0 \\ \alpha_2(q_2-1) & \alpha_2(q_2-1) & -\gamma_2 & p_2+\mu_4 \end{pmatrix}$$

Then we compute the inverse of $\mathcal{V}$ using Wolfram Mathematica:

$$V^{-1} = \begin{pmatrix} \frac{1}{\alpha_2} & 0 & 0 & 0 \\ \frac{p_2(\gamma_2-q_2\mu_2+\mu_2)}{\alpha_2(p_2q_2\mu_2+(\gamma_2+\mu_2)\mu_4)} & \frac{(\gamma_2+\mu_2)(p_2+\mu_4)}{\alpha_2(p_2q_2\mu_2+(\gamma_2+\mu_2)\mu_4)} & \frac{p_2\gamma_2}{p_2q_2\alpha_2\mu_2+\alpha_2\mu_4\mu_2+\alpha_2\gamma_2\mu_4} & \frac{p_2(\gamma_2+\mu_2)}{\alpha_2(p_2q_2\mu_2+(\gamma_2+\mu_2)\mu_4)} \\ \frac{q_2(p_2+\mu_4)}{p_2q_2\mu_2+(\gamma_2+\mu_2)\mu_4} & \frac{q_2(p_2+\mu_4)}{p_2q_2\mu_2+(\gamma_2+\mu_2)\mu_4} & \frac{p_2q_2+\mu_4}{p_2q_2\mu_2+(\gamma_2+\mu_2)\mu_4} & \frac{p_2q_2}{p_2q_2\mu_2+(\gamma_2+\mu_2)\mu_4} \\ \frac{\gamma_2-q_2\mu_2+\mu_2}{p_2q_2\mu_2+(\gamma_2+\mu_2)\mu_4} & \frac{\gamma_2-q_2\mu_2+\mu_2}{p_2q_2\mu_2+(\gamma_2+\mu_2)\mu_4} & \frac{\gamma_2}{p_2q_2\mu_2+(\gamma_2+\mu_2)\mu_4} & \frac{\gamma_2+\mu_2}{p_2q_2\mu_2+(\gamma_2+\mu_2)\mu_4} \end{pmatrix}$$

In order to find R_e we evaluate both matrices F and V^{-1} at our DFE and find the largest eigenvalue of the product of the two matrices

$$R_e = \rho\left(FV^{-1}\right)$$

$$\rho \begin{pmatrix} \frac{I_1\beta_1}{\alpha_2} + \frac{I_1\beta_1p_2(\gamma_2+\mu_2-\mu_2q_2)}{\alpha_2(\mu_4(\gamma_2+\mu_2)+\mu_2p_2q_2)} & \frac{I_1\beta_1(\gamma_2+\mu_2)(\mu_4+p_2)}{\alpha_2(\mu_4(\gamma_2+\mu_2)+\mu_2p_2q_2)} & \frac{I_1\beta_1\gamma_2p_2}{\alpha_2\gamma_2\mu_4+\alpha_2\mu_4\mu_2+\alpha_2\mu_2p_2q_2} & \frac{I_1\beta_1p_2(\gamma_2+\mu_2)}{\alpha_2(\mu_4(\gamma_2+\mu_2)+\mu_2p_2q_2)} \\ \frac{E_1\beta_2}{\alpha_2} + \frac{E_1\beta_2p_2(\gamma_2+\mu_2-\mu_2q_2)}{\alpha_2(\mu_4(\gamma_2+\mu_2)+\mu_2p_2q_2)} & \frac{E_1\beta_2(\gamma_2+\mu_2)(\mu_4+p_2)}{\alpha_2(\mu_4(\gamma_2+\mu_2)+\mu_2p_2q_2)} & \frac{E_1\beta_2\gamma_2p_2}{\alpha_2\gamma_2\mu_4+\alpha_2\mu_4\mu_2+\alpha_2\mu_2p_2q_2} & \frac{E_1\beta_2p_2(\gamma_2+\mu_2)}{\alpha_2(\mu_4(\gamma_2+\mu_2)+\mu_2p_2q_2)} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$R_e = \frac{(\gamma 2+\mu 2)(\mu 4+\mathrm{p2})(\beta 2\mathrm{E1}+\beta 1\mathrm{I1})}{\alpha 2(\mu 4(\gamma 2+\mu 2)+\mu 2\mathrm{p2q2})}$$

This expression can be simplified by dividing whole expression of R_e by the constant part in the numerator

$$= \frac{\beta_2E_1+\beta_1I_1}{\alpha_2(\mu_4(\gamma_2+\mu_2)+\mu_2p_2q_2)}$$

Manipulating the denominator(Den) with some algebra, we obtain a reduced expression of it. We can also apply the reciprocal of some terms and sum them all to get a final expression

$$Den = \alpha_2\left[\frac{\mu_2p_2q_2}{(p_2+\mu_4)(\gamma_2+\mu_2)} + 1 - \frac{p_2}{p_2+\mu_4}\right]$$

$$Den = \alpha_2\left[\frac{p_2q_2}{(p_2+\mu_4)}\left(1-\frac{\gamma_2}{(\gamma_2+\mu_2)}\right) + 1 - \frac{p_2}{p_2+\mu_4}\right]$$

$$Den = \alpha_2\left[\frac{p_2q_2}{(p_2+\mu_4)} - \frac{p_2q_2\gamma_2}{(\gamma_2+\mu_2)(p_2+\mu_4)} + 1 - \frac{p_2}{p_2+\mu_4}\right]$$

$$Den = \alpha_2\left[1 - \frac{p_2}{p_2+\mu_4}\left(1-q_2+\frac{q_2\gamma_2}{\gamma_2+\mu_2}\right)\right]$$

After this, we can put the whole term together to lead up to:

$R_e = (\frac{I_1\beta_1+E_1\beta_2}{\alpha_2})\frac{1}{1-\frac{p_2}{p_2+\mu_4}(1-q_2+\frac{\mu_2}{\gamma_2+\mu_2})}$

A.2 Endemic Equilibrium

In this case, our recidivism model is considered with $\Lambda_2 > 0$. So we arranged the original equations, the 8 equations we have below, in terms of I_2, I_1, E_2, E_1:

1. $\frac{dI_1}{dt} = \Lambda_1 - \beta_1 I_1(I_2+E_2) - \alpha_1 I_1$
2. $\frac{dI_2}{dt} = \Lambda_2 + \beta_1 I_1(I_2+E_2) - \alpha_2 I_2$
3. $\frac{dE_1}{dt} = p_1 O_1 - \beta_2 E_1(I_2+E_2) - \alpha_1 E_1$
4. $\frac{dE_2}{dt} = p_2 O_2 + \beta_2 E_1(I_2+E_2) - \alpha_2 E_2$
5. $\frac{dT_1}{dt} = (I_1+E_1)(q_1\alpha_1) - T_1(\mu_1+\gamma_1)$
6. $\frac{dT_2}{dt} = (I_2+E_2)(q_2\alpha_2) - T_2(\mu_2+\gamma_2)$
7. $\frac{dO_1}{dt} = \gamma_1 T_1 + (1-q_1)(\alpha_1 I_1 + \alpha_1 E_1) - O_1(p_1+\mu_3)$
8. $\frac{dO_2}{dt} = \gamma_2 T_2 + (1-q_2)(\alpha_2 I_2 + \alpha_2 E_2) - O_2(p_2+\mu_4)$

We added e as an exponent to denote the new equations we got for the endemic equilibrium

$$\begin{aligned}
T_1^e &= \frac{q_1\alpha_1}{\mu_1+\gamma_1}(I_1^*+E_1^*)\\
T_2^e &= \frac{q_2\alpha_2}{\mu_2+\gamma_2}(I_2^*+E_2^*)\\
O_1^e &= \frac{\alpha_1}{\mu_3+p_1}[(1-q_1)+q_1\frac{\gamma_1}{\mu_1+\gamma_1}](I_1^e+E_1^e)\\
O_2^e &= \frac{\alpha_2}{\mu_4+p_2}[(1-q_2)+q_2\frac{\gamma_2}{\mu_2+\gamma_2}](I_2^e+E_2^e)\\
\Lambda_1 &= I_1^e(\alpha_1+\beta_1(I_2^e+E_2^e))\\
\Lambda_2 &= \alpha_2 I_2^e - \beta_1 I_1^e(I_2^e+E_2^e)\\
p_1 O_1^e &= E_1^e(\alpha_1+\beta_2(I_2^e+E_2^e))\\
\alpha_2 E_2^e &= p_2 O_2^e + \beta_2 E_1^e(I_2^e+E_2^e)
\end{aligned}$$

Next, we can add the 2 equations for O_1 :

$$\alpha_1\frac{p_1}{\mu_3+p_1}[1-\frac{q_1\mu_1}{\mu_1+\gamma_1}](I_1^e+E_1^e) = E_1^e(\alpha_1+\beta_2(I_2^e+E_2^e)) \tag{9}$$

and calling $r_1 = \frac{p_1}{\mu_3+p_1}[1-\frac{q_1\mu_1}{\mu_1+\gamma_1}]$, to make it simple for later.

Same process for the 2 more for O_2:

$$\alpha_2\frac{p_2}{\mu_4+p_2}[1-\frac{q_2\mu_2}{\mu_2+\gamma_2}](I_2^e+E_2^e) + \beta_2 E_1^e(I_2^e+E_2^e) = \alpha_2 E_2^e \tag{10}$$

with $r_2 = \frac{p_2}{\mu_4+p_2}[1-\frac{q_2\mu_2}{\mu_2+\gamma_2}]$

For the next step we simplify and add O_1 and O_2 and setting the respective r_1 and r_2, we get:

$$\alpha_1 r_1 I_1^e + \alpha_2 r_2 I_2^e = \alpha_1(1-r_1)E_1^e + \alpha_2(1-r_2)E_2^e$$

In order to derive E_1^e :

$$\alpha_1 r_1 I_1^e = \alpha_1(1-r_1)E_1^e + \beta_2 E_1^e(I_2^e + E_2^e)$$
$$E_1^e = \frac{r_1(\Lambda_1 + \Lambda_2 - \alpha_2 I_2^e)}{\alpha_1(1-r_1) + \beta_2(I_2^e + E_2^e)}$$

We also added Λ_1 and Λ_2 in order to get a value for I_1^e

$$I_1^e = \frac{\Lambda_1 + \Lambda_2 - \alpha_2 I_2^e}{\alpha_1}$$

This will help us to find E_2^e,

$$E_2^e = \frac{\alpha_1(\alpha_2 I_2^e - \Lambda_2)}{\beta_1(\Lambda_1 + \Lambda_2 - \alpha_2 I_2^e)} - I_2^e$$

Now, we can plug in E_2^e into E_1^e and add equations 9 and 10 to simplify some terms and end up with this:

$$r_1(\Lambda_1 + \Lambda_2) + (1-r_1)\alpha_2 I_2^e = \frac{r_1(1-r_1)(\Lambda_1 + \Lambda_2 - \alpha_2 I_2^e)^2}{(1-r_1)(\Lambda_1 + \Lambda_2 - \alpha_2 I_2^e) + \beta_2/\beta_1(\alpha_2 I_2^* - \Lambda_2)} + \frac{\alpha_1\alpha_2(1-r_2)(\alpha_2 I_2^e - \Lambda_2)}{\beta_1(\Lambda_1 + \Lambda_2 - \alpha_2 I_2^e)}$$

setting $x = -\Lambda_2 + \alpha_2 I_2^e$ we obtain a cubic equation that can be solve in Wolfram Mathematica numerically, just plugging in the parameters values mentioned in previous sections.

$$(\Lambda_2 + \Lambda_1 r_1 + (1-r_1)x)((1-r_1)(\Lambda_1 - x) + \frac{\beta_2 x}{\beta_1})(\Lambda_1 - x) =$$
$$r_1(1-r_1)(\Lambda_1 - x)^3 + \frac{\alpha_1\alpha_2}{\beta_1}(1-r_2)x((1-r_1)(\Lambda_1 - x) + (\beta_2/\beta_1)x)$$

Using Mathematica to solve this cubic equation will lead you to 3 values for I_2^e but the one that works for this study is just when $I_2^e = 19236.9$ because, it's the value that give us all positive numbers whenever we replace it on the other equations expressed in terms of I_2. The following values for the other variables of the endemic equilibrium after replacing the parameters values are:

$$I_1^e = \frac{-\alpha_2\left(I_2\right) + \lambda_1 + \lambda_2}{\alpha_1} = 2088.1$$
$$E_2^e = \frac{\alpha_1\left(\alpha_2 I_2 - \lambda_2\right)}{\beta_1\left(-\alpha_2\left(I_2\right) + \lambda_1 + \lambda_2\right)} - I_2 = 89$$
$$E_1^e = \frac{r_1\left(-\alpha_2\left(I_2\right) + \lambda_1 + \lambda_2\right)}{\beta_2\left(E_2 + I_2\right) + \alpha_1\left(1 - r_1\right)} = 704.4$$
$$T_1^e = \frac{\left(E_1 + I_1\right)\left(\alpha_1 q_1\right)}{\gamma_1 + \mu_1} = 98$$
$$T_2^e = \frac{\left(E_2 + I_2\right)\left(\alpha_2 q_2\right)}{\gamma_2 + \mu_2} = 372$$
$$O_1^e = \frac{\alpha_1 E_1 - \beta_2 E_1\left(E_2 + I_2\right)}{p_1} = 1,050$$
$$O_2^e = \frac{\alpha_2 e_2 - \beta_2 e_1\left(e_2 + i_2\right)}{p_2} = 2,149$$

These values are the ones which allow us to construct the cost analysis.

References

[1] Wilson E Alvarez, Itelhomme Fene, Kimberly Gutstein, Brenda Martinez, Diego Chowell5 Anuj Mubayi, and Luis Melara. Prisoner reform programs, and their impact on recidivism. Technical report, Technical report, MTBI–09–06M, 2012.

[2] Bechir Amdouni, Marlio Paredes, Christopher Kribs, and Anuj Mubayi. Why do students quit school? implications from a dynamical modelling study. *Proc. R. Soc. A*, 473(2197):20160204, 2017.

[3] The World Bank. Population, total. Technical report, The World Bank, 2018. `https://data.worldbank.org/indicator/SP.POP.TOTL?end=2017&start=1960`.

[4] Budget Bureau of Planning and Research. Fy 2017 operating per capita cost report. *ARIZONA DEPARTMENT OF CORRECTIONS*, pages 18–20, 2017. `https://corrections.az.gov/sites/default/files/REPORTS/Operating_Per_Capita/adc-percapcostreport_fy2017-final-sectvmod.pdf`.

[5] Lois M Davis, Robert Bozick, Jennifer L Steele, Jessica Saunders, and Jeremy NV Miles. *Evaluating the effectiveness of correctional education: A meta-analysis of programs that provide education to incarcerated adults*. Rand Corporation, 2013.

[6] Marco Hamins-Puértolas, Jessica Conrad, Genesis Islas, Adrien Bossogo-Egoume, Christopher Kribs, and Benjamin Morin. Minimizing recidivism by optimizing profit: a theoretical case study of incentivized reform in a louisiana prison. 2015.

[7] James Holland Jones. Notes on r0. *Califonia: Department of Anthropological Sciences*, 2007.

[8] Christopher Kribs, Anuj Mubayi, Maia Martcheva, and Carlos Castillo-Chávez. A cost-based comparison of quarantine strategies for new emerging diseases. 2010.

[9] Xia Ma, Yicang Zhou, and Hui Cao. Global stability of the endemic equilibrium of a discrete sir epidemic model. *Advances in Difference Equations*, 2013(1):42, 2013.

[10] Arizona Department of Corrections. Arizona inmate recidivism study. Technical report, Arizona Department of Corrections, 2005. `https://aci.az.gov/images/Recidivism/ArizonaInmateRecidivismStudy.pdf`.

[11] Arizona Department of Corrections. Corrections at a glance. Technical report, Arizona Department of Corrections, 2017. `https://corrections.az.gov/reports-documents/reports/corrections-glance`.

[12] Arizona Department of Corrections. Inmate education. Technical report, 2018. `https://corrections.az.gov/workforce-development#work%20based%20education`.

[13] Arizona Department of Corrections. Overview and budget information. Technical report, Arizona Department of Corrections, 2018. `https://www.azleg.gov/jlbc/19adcagypres.pdf`.

[14] California Department of Corrections and Rehabilitation. Evaluation of preventing parolee failure program. Technical report, 1997. `https://www.ncjrs.gov/pdffiles1/Digitization/180542NCJRS.pdf`.

[15] Florida Department of Corrections. Inmate population: Tabe (tests of adult basic education). Technical report, Florida Department of Corrections, 2016. `http://www.dc.state.fl.us/pub/annual/1516/stats/ip_grade_level.html`.

[16] Florida Department of Corrections. Annual report florida department of corrections. Technical report, Corrections Foundation, 2017. `http://www.dc.state.fl.us/pub/annual/1617/FDC_AR2016-17.pdf`.

[17] US House of Representatives. Comprehensive crime control act of 1984. Act S.1762, United States House of Representatives, February 1984. `https://www.ncjrs.gov/pdffiles1/Digitization/123365NCJRS.pdf`.

[18] Carles L Ryan. Adc fiscal year 2019 decision package summary. Technical report, 2017. https://corrections.az.gov/sites/default/files/documents/BudgetReq/adc-fy-2019-operating-budget-request_090117.pdf.

[19] Charles L. Ryan. Strategic plan fy2018-2022. Technical report, Arizona Department of Corrections, 2017.
https://corrections.az.gov/sites/default/files/documents/PDFs/adc-5-yr-stratplan-fy2018-fy2022_rev0_final_122017.pdf.

[20] Charles L. Ryan. Five-year strategic plan fy 2018 to fy 2022. Technical report, Arizona Department of Corrections, 2018.
https://corrections.az.gov/sites/default/files/documents/PDFs/adc-5-yr-stratplan-fy2018-fy2022_final-010617.pdf.

[21] Arizona State University. Standard cost of attendance. Technical report, Arizona State University, 2018.
https://students.asu.edu/standard-cost-attendance.

[22] Pauline Van den Driessche and James Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical biosciences*, 180(1-2):29–48, 2002.

[23] Roy Walmsley. World prison population list. Technical Report 11th Edition, Institute for Criminal Policy Research, 2015.
http://www.prisonstudies.org/sites/default/files/resources/downloads/world_prison_population_list_11th_edition_0.pdf.

Controlling Foodborne Infections in Lettuce: Testing and Cleaning Methods for Curbing the Spread of *E. coli* O157:H7

Emily Dorn[1], Zakiya Jones[2], Crystal Mackey[3], Jessica C. Reyes[4], Mugdha Thakur[5], Victor M. Moreno[5], Carlos Castillo-Chavez[5], Carlos Castillo-Garsow[6] and Baojun Song[7]

[1]Olivet College
[2]Pomona College
[3]University of Texas at Arlington
[4]Kennesaw State University
[5]Arizona State University
[6]Eastern Washington University
[7]Montclair State University

Abstract

E. coli O157:H7 bacteria tends to contaminate leafy green vegetable farms regularly, therefore promoting a recurrence in foodborne disease outbreaks. In previous studies, *E. coli* in water has been a focus. However, the farming industry is just now understanding the danger of *E. coli* in soil. In spring 2018, lettuce farms in the Yuma region contracted the *E. coli* bacteria and caused 210 human cases nationwide. As a result of this study, a new mathematical framework is proposed to capture the dynamics of the spread of *E. coli* in lettuce due to contaminated soil and equipment. In particular, this framework explores the impact of soil treatment and equipment sanitation since they are essential to the growing process of lettuce.

1 Introduction

The name *Escherichia coli* encompasses a large group of bacteria. Most strains of this bacteria are harmless to humans, but strains such as the Shiga toxin-producing *E. coli* (STEC) O157:H7 are pathogenic to humans [8]. In this study, mention of *E. coli* equates to the STEC O157:H7 strain. *E. coli* spreads to humans through contaminated food, usually ground beef and produce [9]. The Centers for Disease Control and Prevention (CDC) estimate that each year 96,534 individuals are infected with the strain STEC O157:H7 [25] and that 46% of outbreaks in the United States occur because of contaminated produce [5]. Leafy greens are the most common produce that are contaminated with *E. coli* [7], which is supported by the *E. coli* outbreaks from 2011, 2012, 2013, 2017, and 2018 [6]. The 2018 *E. coli* outbreak in romaine lettuce was linked to farms in the Yuma region. The commercial lettuce farms in Yuma, Arizona are responsible for 90% of the nations leafy greens [3] during the winter. As a result, this outbreak had wide reaching consequences. As of June 28, 2018, there were 36 states affected by the Yuma lettuce outbreak as seen in Figure 1. In this study, we focus specifically on the dynamics of the spread of *E. coli* in lettuce.

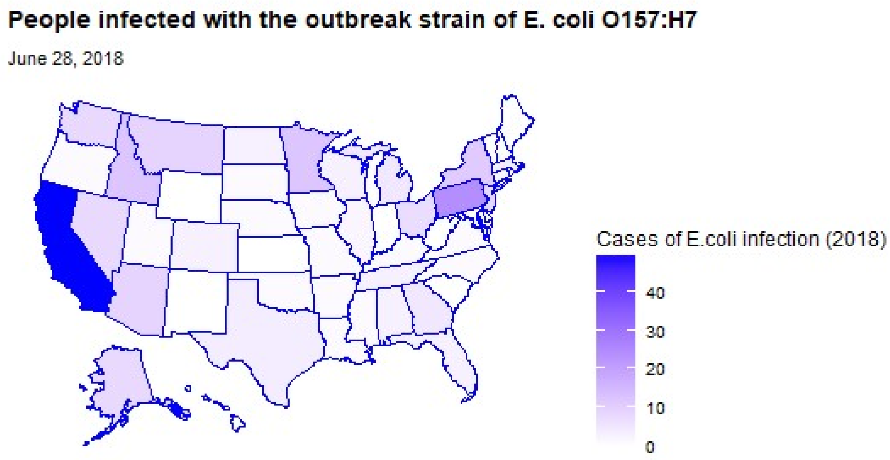

Figure 1: In this United States case count map of *E. coli*, a total of 210 human cases across 36 states were linked to romaine lettuce [18].

Although STEC O157:H7 is not the most common foodborne illness, it causes some of the most severe symptoms. When people consume *E. coli* contaminated food, they become infected and can experience severe symptoms such as bloody diarrhea, abdominal cramping, and vomiting [9]. Approximately, five to ten percent of *E. coli* cases lead to hemolytic-uremic syndrome (HUS) which can cause kidney failure and death [9]. Furthermore, these numbers do not encompass the full impact of the disease as cases are not reported when the illness is not linked to *E. coli* or the person does not go to the doctor. Only 20% of illnesses due to a foodborne outbreak are reported [8]. While the lettuce, itself, is unharmed by the presence of STEC O157:H7 [26], it can transmit *E. coli* to humans.

E. coli lives in the lower intestines of humans and animals as a part of the digestive system [23]. Cattle are the main source of contamination [10]. Infected cattle can contaminate their feces with *E. coli* from their intestines; this is called shedding. Cows can shed as many as 10^6 to 10^9 colony forming units (CFU) per gram of feces [16]. The waste from the infected cow is processed to use as manure for growing crops including lettuce. If this waste is improperly composted, it will still contain *E. coli* which will spread to the soil when manure is applied [17]. Consequently, this contaminated manure infects the lettuce that is being cultivated. Unknowingly, the *E. coli*-infected lettuce is grown, harvested, processed, and shipped out to stores for consumers to eat and become ill [17]. Farming equipment is an important mechanism of transmission [11]. Agricultural equipment enters the farming process at different times during the life cycle of the lettuce; they are used for the laying of manure, pre-planting ground preparation, planting, thinning, and harvesting [21]. Farm vehicles used in the growing season include the chisel plow, disk harrow, stanhay type planter, rotary spiker, and harvesting vehicle [21]. In addition, farm tools like long knives are used to cut heads of lettuce during harvesting. During these times, if equipment is used on contaminated soil or used to manipulate contaminated manure, it can transmit *E. coli* to uninfected soil and spread the infection.

There are several points of intervention throughout this process. Lettuce can be sanitized and disinfected within the normal process of preparing lettuce for consumers.

However, this will not be effective if the lettuce has internalized the bacteria. Internalization occurs when a lettuce seed is planted in contaminated soil and uptakes *E. coli* in its roots as it grows [26]. If this happens, then, it is nearly impossible to remove the bacteria from the lettuce. As a result, other points of intervention that prevent *E. coli* from coming into contact with soil at any point must be taken into account to limit internalization. One way to do so would be to prevent equipment from touching contaminated soil and manure.

Previous studies have shown that prevention during the preharvest process is crucial for reducing contamination. Past work by Franz, *et. al.* (2008) focused on the ecological factors that lead to the growth of STEC O157:H7 in lettuce. The probabilistic model concentrated on manure-amended soil through the production process. This study considered variables, such as herd density, manure storage time intervals, and manure quantity in order to estimate the probability of *E. coli* infected lettuce. The results show manure and soil management to be influential in preventing pathogenic *E. coli* in lettuce. Furthermore, the study states that there is a high correlation between the initial prevalence of contaminated manure and the probability of contaminated lettuce [17].

Our research focuses on testing and treatment methods to identify and treat contaminated soil and on sanitation methods of contaminated farm equipment through mathematical modeling. Some of the current procedures and standards of soil testing in a commercial farm setting include sending in soil samples to test in labs and at home soil testing kits [20]. If STEC O157:H7 is discovered in soil, then the soil is treated by no longer planting lettuce and letting the soil dry out from the sun via UV rays [20]. The standard for cleaning farm equipment consists of four stages. The first stage involves washing to loosen the soil on the surface. The second stage incorporates the use of detergent and scrubbing to break up the adhesion of the microorganisms. The next stage is rinsing that removes the loosened soil and detergent. The last stage is applying sanitizer to kill as much microorganisms as possible [27]. The frequency of sanitation is at the discretion of the farm, but the Food and Drug Administration (FDA) recommends that each farm develops their own sanitation standard operating procedures and schedules for sanitation [14].

As a result, we develop and analyze a mathematical model that describes the interaction of infected manure, contaminated soil, farm equipment, and lettuce. Our analysis uses sensitivity analysis of the basic reproduction number to determine the effect of treatment and sanitation on the yield of healthy lettuce.

2 Methods and Model

In our model, we first consider four state variables. The first two are clean equipment (E_c) and contaminated equipment (E_I). The other two state variables are healthy lettuce (L_S) and contaminated lettuce (L_I). The dynamics between the state variables are described in the following paragraphs.

We then assume a proportion of contaminated manure, ρ,is applied to the soil for fertilization. Consequently, we assume that *E. Coli* then colonizes the rest of the soil as a rate, r. In our model, the equation for the change in the proportion of infected soil, P, over time, t, is based off the Levins Model. The Levins model was developed to implement migration and extinction of a population in patches by utilizing a logistic growth equation [22]. We modified the Levins model to incorporate $M\rho$ into our model to reflect the proportion of infected manure that contaminates a proportion of soil at the start of cultivation. We assume that the infected manure infects the soil at the following rate,

$$rM\rho P(1 - P).$$

Additionally, soil is treated at a rate $t_A P$ and the *E. coli* in the contaminated soil naturally dies out at rate dP. Manure and soil are essential in the growing process of lettuce.

The change in P over time t is shown below,

$$\frac{dP}{dt} = rM\rho P(1-P) - (d+t_A)P.$$

The germination of healthy lettuce plants occurs at a rate, αM. Germination occurs when a seed sprouts under favorable conditions such as appropriate water intake and temperature. We assume lettuce becomes contaminated in two ways. When the lettuce seed or plant comes into contact with infected soil, *E. coli* transmits to the lettuce at a rate of β_{L_S}. As a result, lettuce can internalize the *E. coli* from the soil and then the seed or plant becomes contaminated. In addition, healthy lettuce becomes infected when it comes in contact with infected equipment and the bacteria transmits at a rate $\hat{\beta}_{L_S}$. The per capita death rate of lettuce is μ_L which occurs at different stages of the growing season like during the thinning and harvesting process [21].

We consider farm equipment as our vector for *E. coli* spreading during the farming process. It is further assumed that new equipment such as tools and farm vehicles are acquired at a rate, Λ_E. Equipment moves from the clean, E_c, to contaminated, E_I, compartment when it comes in contact with *E. coli* contaminated manure and transmits *E. Coli* at a rate, β_M. Likewise, the term $\beta_{E_c} P \frac{E_c}{N_E}$, shows how clean equipment is contaminated as a result of its interaction with contaminated soil per unit time. The contact between equipment and soil occurs during the preplanting, thinning, fertilization, and harvesting stages. We consider all of the equipment and tools mentioned in Section 1 as the total population of equipment, N_E. The cleaning rate, γ_c, describes the rate at which a farmer randomly cleans infected equipment without knowing which ones are contaminated and moves equipment from the E_I to E_c compartment given that the farmer can only clean a certain amount of equipment per unit time, γ_c. However, if contaminated equipment cannot return to a clean state then it is discarded from the system at a rate, δ_E. Additionally, clean and contaminated equipment are considered to have gone through a per capita disposal rate, μ_E, when they are no longer usable. For example, dulled or broken knives that are discarded during harvest. Furthermore, homogeneous mixing of equipment and lettuce is assumed.

All of the dynamics are described in equations (1 – 4) and represented in Figure 2.

$$\frac{dP}{dt} = rM\rho P(1-P) - (d+t_A)P, \tag{1}$$

$$\frac{dE_c}{dt} = \Lambda_E + \gamma_c \frac{E_I}{N_E} - \beta_{E_c} P \frac{E_c}{N_E} - \beta_M M \rho E_c - \mu_E E_c, \tag{2}$$

$$\frac{dE_I}{dt} = \beta_{E_c} P \frac{E_c}{N_E} + \beta_M M \rho E_c - \gamma_c \frac{E_I}{N_E} - (\mu_E + \delta_E) E_I, \tag{3}$$

$$\frac{dL_I}{dt} = \beta_{L_S} P L_S + \hat{\beta}_{L_S} \frac{E_I}{N_E} L_S - \mu_L L_I, \tag{4}$$

$$\frac{dL_S}{dt} = \alpha M - \beta_{L_S} P L_S - \hat{\beta}_{L_S} \frac{E_I}{N_E} L_S - \mu_L L_S, \tag{5}$$

where $N_E = E_c + E_I$.

Because the total population of lettuce, N_L is assumed to approach $\frac{\alpha M}{\mu_L} = k$, the lettuce population is asymptotically constant.

From equation (1), we can see that P is between 0 and 1 because $\frac{dP}{dt}\big|_{P=0} = 0$, and $\frac{dP}{dt}\big|_{P=1} < 0$. From equation (2) and (3), we can see that $0 < E_c + E_I \leq \frac{\Lambda_E}{\mu_E} = k_E$

(carrying capacity of equipment). So the triangle region $\Delta = \{0 \leq E_I + E_c \leq k_E, E_I \geq 0, E_c \geq 0\}$ is positive invariant. Similarly, from equation (5 – 4), we know $0 \leq L_I \leq k$. Therefore, the domain of interest of our model is

$$\Omega = [0, 1] \times \Delta \times [0, k]$$

which is a positive invariant for the system (1 – 4). We will analyze the model within this domain.

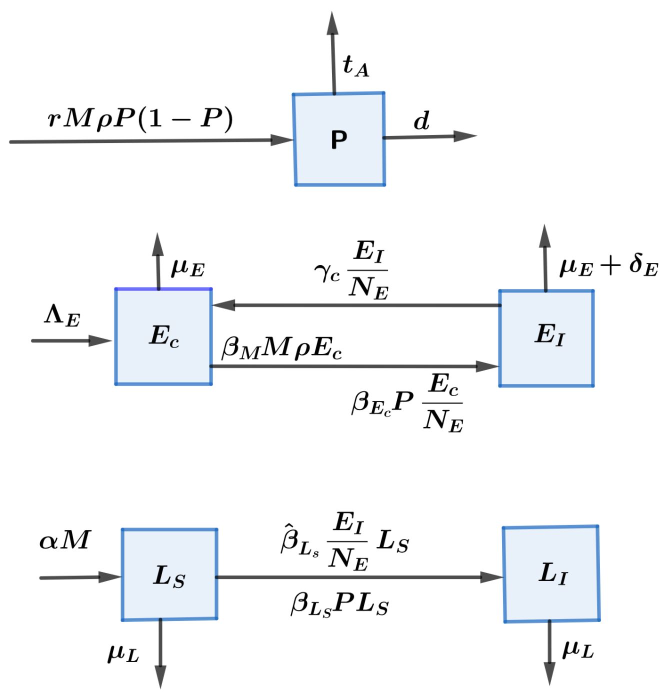

Figure 2: Flow diagram describing the interactions from equations (1 - 4). P is the proportion of infected soil, E_c clean equipment, E_I contaminated equipment, L_S healthy lettuce, and L_I infected lettuce.

3 Parameters

The first assumption we make for our parameters is that we define the size of a field to be 1 km^2. The Food Safety Modernization Act states that if a farm's water is known to be infected with *E. coli*, then the water is to be tested four times within a growing season [15]. We assume that this is the same for soil and calculated our treatment rate by dividing the number of times soil is tested by the number of days in a growing season, 152 days [21]. The parameter, δ_E, was calculated based on a strategy of some farmers for replacing their equipment. This particular strategy consists in replacing one or two pieces of equipment every year [19]; consequently, we divided the 2 pieces of equipment by the total number of days in a year to get, $\frac{2}{365} \approx 0.005$. Λ_E, the equipment acquired per day, is estimated to be some amount of equipment, less than 25. μ_E has to be less than Λ_E, because when μ_E is computed it is multiplied by the total population of equipment. The cleaning rate of equipment, γ_c, was calculated based on the presumption that farmers clean 10 pieces of equipment weekly, biweekly, monthly, or never, which is shown by dividing 1 by the number of days between the cleaning, $\frac{10}{7}$, $\frac{10}{14}$, $\frac{10}{30}$, and 0 respectively. Therefore, the range of γ_c is between 0 and $\frac{10}{7}$.

Between 25,500 to 41,500 seedlings of lettuce are planted at the start of the lettuce growing season per acre [28], which we use to calculate the germinating rate of healthy lettuce, α, by first dividing by 0.004 to convert it to the number of lettuce per km^2, 6,375,000 to 10,375,000. Then, we divide 6,375,000 and 10,375,000 by the amount of manure times the number of days in our growing season, $\frac{6375000}{2300000*152} \approx 0.0182$ and $\frac{10375000}{2300000*152} \approx 0.03$. We can also use these numbers to compute the per capita death rate by multiplying the amount of manure by α and then dividing by the number of lettuce per km^2, $\frac{2300000*0.0182}{6375000} \approx 0.0066$ to $\frac{2300000*0.03}{10375000} \approx 0.0067$. All of the β values can be varied. The parameters, ρ and r are assumed to be proportions, 0 to 1. Parameter values M and d were found from other sources [17], [1] and the only computations involving these parameters are conversions to be consistent with the units used in this study.

Symbol	**Parameters**	**Estimations**	**Reference**
r	Growth rate of infected soil ($\frac{1}{kg \times day}$)	[0,1]	Estimate
M	Quantity of manure (kg)	$2,300,000$	[17]
ρ	Proportion of infected manure	$[0,1]$	Estimate
t_A	Treatment rate of soil as a result of testing ($\frac{1}{day}$)	Varies	Estimate
d	Per capita death rate of *E. coli* in soil ($\frac{1}{\text{day}}$)	0.3326	[1]
Λ_E	Rate equipment is acquired ($\frac{equipment}{day}$)	0.164	[19]
α	Germinating rate of healthy lettuce ($\frac{lettuce}{day \times mass}$)	$[0.0182, 0.03]$	[28]
β_{E_c}	Rate of transmission per equipment ($\frac{1}{equipment \times day}$)	Varies	Estimate
β_M	Transmission rate of manure ($\frac{1}{mass \times day}$)	Varies	Estimate
β_{L_S}	Contamination rate of healthy lettuce due to infected soil ($\frac{1}{day}$)	Varies	Estimate
$\hat{\beta}_{L_S}$	Contamination rate of lettuce due to infected equipment ($\frac{1}{day}$)	Varies	Estimate
γ_c	Cleaning rate of equipment ($\frac{equipment}{day}$)	Varies	Estimate
μ_E	Per capita disposal rate of equipment ($\frac{1}{day}$)	0.164	[19]
μ_L	Per capita death rate of lettuce ($\frac{1}{day}$)	0.0066	[28]
δ_E	Removal rate of infected equipment due to inability to clean ($\frac{1}{day}$)	Varies	Estimate

Table 1: Symbols, Definitions, and Parameter Estimates.

4 Analysis

In order to facilitate our mathematical analysis, we begin by re-scaling our model. Then, we find equilibria for the system and determine the conditions for their existence and stability as well as determining the important threshold, R_0^S expression. Finally, we examine the global dynamics for our system.

4.1 Re-scaled Model

Since $L_I + L_S = k$, we reduce the system (1 – 4) in terms of the following system of four equations

$$\frac{dP}{dt} = rM\rho P(1-P) - (d+t_A)P, \tag{6}$$

$$\frac{dE_c}{dt} = \Lambda_E + \gamma_c \frac{E_I}{N_E} - \beta_{E_c} P \frac{E_c}{N_E} - \beta_M M\rho E_c - \mu_E E_c, \tag{7}$$

$$\frac{dE_I}{dt} = \beta_{E_c} P \frac{E_c}{N_E} + \beta_M M\rho E_c - \gamma_c \frac{E_I}{N_E} - (\mu_E + \delta_E)E_I, \tag{8}$$

$$\frac{dL_I}{dt} = \beta_{L_S} P(k - L_I) + \hat{\beta}_{L_S} \frac{E_I}{N_E}(k - L_I) - \mu_L L_I, \tag{9}$$

where $N_E = E_c + E_I$ and $k = \frac{\alpha M}{\mu_L}$

The equations (6 – 9) are re-scaled using the following equivalences.

$$t = \frac{\tau}{\rho M r},\ P = \frac{\gamma_c}{\beta_{E_c}}x,\ E_c = \frac{\gamma_c}{\rho M r}y,\ E_I = \frac{\gamma_c}{\rho M r}z, \text{ and } L_I = kw.$$

The nondimensionalized system of (6 – 9) becomes

$$\frac{dx}{d\tau} = x(1 - Ax) - G_0 x, \tag{10}$$

$$\frac{dy}{d\tau} = G_1 + \frac{z}{y+z} - \frac{xy}{y+z} - G_2 y, \tag{11}$$

$$\frac{dz}{d\tau} = \frac{xy}{y+z} + G_4 y - \frac{z}{y+z} - G_6 z, \tag{12}$$

$$\frac{dw}{d\tau} = G_7 x(1-w) + G_8 \frac{z}{y+z}(1-w) - G_9 w, \tag{13}$$

where, $A = \frac{\gamma_c}{\beta_{E_c}}$, $G_0 = \frac{d+t_A}{\rho M r}$, $G_1 = \frac{\Lambda_E}{\gamma_c}$, $G_2 = \frac{\beta_M M\rho + \mu_E}{\rho M r}$, $G_4 = \frac{\beta_M}{r}$, $G_6 = \frac{\mu_E + \delta_E}{\rho M r}$, $G_7 = \frac{\beta_{L_S}\gamma_c}{\beta_{E_c}\rho M r}$, $G_8 = \frac{\hat{\beta}_{L_S}}{\rho M r}$, and $G_9 = \frac{\mu_L}{\rho M r}$.

The domain of interest, Ω, is changed to $\Omega^{'} = [0, 1/A] \times \Delta^{'} \times [0,1]$ where $\Delta^{'} = \{0 \leq x + y \leq \frac{k_E M r \rho}{\gamma_c}\}$.

4.2 Contamination-Free Equilibrium of Soil

We will now do a full analysis of our re-scaled model (10–13). Firstly, the contamination-equilibrium of soil is ($x^* = 0$, y^*, z^*, w^*) where,

$$
\begin{aligned}
y^* &= \frac{-(1+2G_1)G_2 + G_4 + G_1(G_4+G_6)}{2G_2(G_4+G_6-G_2)} \\
&+ \frac{\sqrt{G_2^2 - 2G_2(G_4+G_1G_4-G_1G_6) + (G_4+G_1G_4+G_1G_6)^2}}{2G_2(G_4+G_6-G_2)} \\
z^* &= \frac{G_1 + (G_4-G_2)y^*}{G_6}, \\
w^* &= \frac{G_8 z^*}{G_9(y^*+z^*) + G_8 z^*}.
\end{aligned}
$$

Given that y^* exists, z^* exists if $y^* < \frac{\Lambda_E M r \rho}{\gamma_c \mu_E}$. Also, w^* exists if both y^* and z^* exist. To determine the stability of the contamination-free soil equilibrium, we find the Jacobian matrix

$$
\begin{pmatrix}
-G_0+1 & 0 & 0 & 0 \\
-\frac{y^*}{y^*+z^*} & -G_2 - \frac{z^*}{(y^*+z^*)^2} & -\frac{y^*}{(y^*+z^*)^2} & 0 \\
\frac{y^*}{y^*+z^*} & G_4 + \frac{z^*}{(y^*+z^*)^2} & -G_6 - \frac{y^*}{(y^*+z^*)^2} & 0 \\
G_7(1-w^*) & \frac{G_8(-1+w^*)z^*}{(y^*+z^*)^2} & -\frac{G_8(-1+w^*)z^*}{(y^*+z^*)^2} & -G_9 - \frac{G_8 z^*}{y^*+z^*}
\end{pmatrix}
$$

and, solve for the eigenvalues,

$$
\lambda_1 = 1 - G_0 \text{ and } \lambda_2 = -G_9 - \frac{G_8 z^*}{y^*+z^*}.
$$

After reducing the Jacobian to a 2×2 matrix:

$$
\begin{pmatrix}
-G_2 - \frac{z^*}{(y^*+z^*)^2} & -\frac{y^*}{(y^*+z^*)^2} \\
G_4 + \frac{z^*}{(y^*+z^*)^2} & -G_6 - \frac{y^*}{(y^*+z^*)^2}
\end{pmatrix}.
$$

In order to determine stability, we want to determine the conditions for when λ_1 and λ_2 are negative. In particular, $-G_0+1$ is negative when $G_0 > 1$ or $\frac{t_A+d}{Mr\rho} > 1$, and $-G_9 - \frac{G_8 z^*}{y^*+z^*}$ is always negative. To determine when the other two eigenvalues have negative real parts we check when the trace of the 2×2 matrix is negative and the determinant is positive. Obviously, the trace, $-G_2 - \frac{z^*}{(y^*+z^*)^2} - G_6 - \frac{y^*}{(y^*+z^*)^2}$ is always negative. The determinant

$$
\frac{G_2 y^* - G_4 y^* + G_6 z^* + G_2 G_6 (y^*+z^*)^2}{(y^*+z^*)^2},
$$

is positive because $G_4 < G_2$ since $\frac{\beta_M}{r} < \frac{\beta_M M\rho + \mu_E}{\rho M r}$. Therefore, the equilibrium is locally asymptotically stable when $G_0 > 1$. We collect the above analysis into the following theorem.

Theorem 1. *The contamination-free equilibrium of soil of system (10-13) is asymptotically stable when* $R_0^S = \frac{1}{G_0} = \frac{rM\rho}{t_A+d} < 1$.

Definition 1. R_0^S is the basic reproductive number for soil.

4.3 Endemic Equilibrium of Soil

The other equilibrium for equation (10) is $x_2^* = \frac{1}{A}(1 - G_0)$ which exists if $R_0^S > 1$. Let $(x_2^* = \frac{1}{A}(1 - G_0), y_2^*, z_2^*, w_2^*)$ be the endemic equilibrium of soil. Here, y^* is given by the following quadratic equation

$$ay^2 + by + c = 0 \tag{14}$$

where,

$$\begin{aligned}
a &= G_2(G_2 - G_4 - G_6) = -\frac{\delta_E(\beta_M M\rho + \mu_E)}{\rho M r}, \\
b &= G_6\left(G_1 - \frac{1}{A} + \frac{G_0}{A}\right) + G_4(G_1 + 1) - G_2 - 2G_1G_2 = \\
&\frac{\mu_E(\Lambda_E - \beta_{E_c} + \gamma_c - 2\Lambda_E\beta_M M\rho r) + \delta_E(-\beta_{E_c} + \Lambda_E) + \rho\beta_M\Lambda_E M}{\rho M r\gamma_c} \\
&+ (\mu_E + \delta_E)(d + t_A)(\beta_{E_c})(\rho M r)^2\gamma_c, \\
c &= G_1(G_1 + 1) = \frac{\Lambda_E}{\gamma_c}\left(\frac{\Lambda_E}{\gamma_c} + 1\right).
\end{aligned}$$

We know that $a < 0$ and $c > 0$, then the product of the two roots is less than zero. Therefore, equation (14) has one positive and one negative root and the positive root is the endemic solution. Then y_2^* is a positive root. We get

$$\begin{aligned}
y_2^* = &\frac{\sqrt{-4(-G_1 - G_1^2)(G_2^2 - G_2G_4 + G_2G_6) + (-G_1G_6 - \frac{G_0G_6 - G_6}{A} - G_4G_1 - G_4 + G_2 + 2G_1G_2)^2}}{2(G_2^2 - G_2G_4 + G_2G_6)} \\
&+ \frac{-G_1G_6 - \frac{G_0G_6 - G_6}{A} - G_4G_1 - G_4 + G_2 + 2G_1G_2}{2(G_2^2 - G_2G_4 + G_2G_6)}
\end{aligned}$$

and direct computation gives us

$$\begin{aligned}
z_2^* &= \frac{G_1 + (G_4 - G_2)y_2^*}{G_6}, \\
w_2^* &= \frac{\frac{(1-G_0)G_7y}{A} + \frac{(1-G_0)G_7z}{A} + G_8z}{\frac{(1-G_0)G_7y}{A} + G_9y + \frac{(1-G_0)G_7z}{A} + G_8z + G_9z}.
\end{aligned}$$

Since there exists a positive y_2^*, z_2^* is positive when the positive y_2^* is plugged in. In order for $z_2^* > 0$, the condition for existence $y_2^* < \frac{\Lambda_E M r\rho}{\gamma_c\mu_E}$ and $(x_2^*, y_2^*, z_2^*, w_2^*)$ exists when $R_0^S > 1$.

Similarly to the contamination-free equilibrium of soil stability, we can use a Jacobian to determine the stability of the endemic equilibrium corresponding to $x_2^* = \frac{1}{A}(1 - G_0)$

$$\begin{pmatrix} -1+G_0 & 0 & 0 & 0 \\ -\frac{y_2^*}{y_2^*+z_2^*} & \frac{(-1+G_0)z_2^*-A(z_2^*+G_2(y_2^*+z_2^*)^2)}{A(y_2^*+z_2^*)^2} & \frac{(1+A-G_0)y_2^*}{A(y_2^*+z_2^*)^2} & 0 \\ \frac{y_2^*}{y_2^*+z_2^*} & \frac{z_2^*-G_0z_2^*+A(z_2^*+G_4(y_2^*+z_2^*)^2)}{A(y_2^*+z_2^*)^2} & -G_6+\frac{(-1-A+G_0)y_2^*}{A(y_2^*+z_2^*)^2} & 0 \\ G_7(1-w_2^*) & -\frac{G_8(1-w_2^*)z_2^*}{(y_2^*+z_2^*)^2} & -\frac{G_8(-1+w_2^*)y_2^*}{y_2^*+z_2^*)^2} & \frac{(-1+G_0)G_7}{A}-G_9-\frac{G_8z_2^*}{y_2^*+z_2^*} \end{pmatrix}.$$

Two eigenvalues of this 4×4 matrix are

$$\lambda_1 = -1+G_0 \qquad \text{and} \qquad \lambda_2 = \frac{(-1+G_0)G_7}{A} - G_9 - \frac{G_8 z_2^*}{y_2^*+z_2^*}.$$

The eigenvalues λ_1 and λ_2 are negative with the condition that $G_0 < 1$. Similar to the contamination-free equilibrium of soil, we can analyze a simplified 2×2 version of the Jacobian matrix

$$J = \begin{pmatrix} \frac{(-1+G_0)z_2^*-A(z_2^*+G_2(y_2^*+z_2^*)^2)}{A(y_2^*+z_2^*)^2} & \frac{(1+A-G_0)y_2^*}{A(y_2^*+z_2^*)^2} \\ \frac{z_2^*-G_0z_2^*+A(z_2^*+G_4(y_2^*+z_2^*)^2)}{A(y_2^*+z_2^*)^2} & -G_6+\frac{(-1-A+G_0)y_2^*}{A(y_2^*+z_2^*)^2} \end{pmatrix}.$$

Since the two outermost eigenvalues are negative, we only need to show that the trace of the 2×2 matrix is negative and the determinant is positive for the endemic equilibrium to be stable

$$Trace(J) = \frac{(-1+G_0)z_2^* - A(z_2^*+G_2(y_2^*+z_2^*)^2)}{A(y_2^*+z_2^*)^2} - G_6 + \frac{(-1-A+G_0)y_2^*}{A(y_2^*+z_2^*)^2} < 0$$

when $G_0 < 1$. The determinant is

$$Det(J) = \frac{y_2^*((1+A-G_0)(G_2-G_4)+AG_2G_6y_2^*)+G_6(1+A-G_0+2AG_2y_2^*)z_2^*+AG_2G_6z_2^*}{A(y_2^*+z_2^*)^2}$$

Since we have already set the condition that $G_0 < 1$ and $G_4 < G_2$, the determinant is positive. Based on the condition that $R_0^S > 1$ then $Trace(J) < 0$ and $Det(J) > 0$. Therefore, the endemic equilibrium is stable.

Theorem 2. *The endemic equilibrium of soil of system (10-13) is asymptotically stable when* $R_0^S = \frac{1}{G_0} = \frac{rM\rho}{t_A+d} > 1$.

4.4 Global Dynamics

The contamination-free equilibrium of soil, $x^* = 0$ and the endemic equilibrium of soil $x_2^* = \frac{1}{A}(1-G_0)$ are globally stable when $R_0^S < 1$ and $R_0^S > 1$ respectively. By the limiting equation theorem [4], we can substitute x^* and x_2^* into equations (11 – 12). Equations (11 – 12) are a closed planar system. If we can rule out that our system has closed trajectories in our domain of interest, Ω', then our local equilibrium stability becomes globally stable. We can prove this by using the Dulac Criterion [2, 172].

Theorem 3. *Dulac's Criterion: If $D(y,z)$ in C^1 in a region $B \subseteq \mathbb{R}^2$ (simply connected) and $\frac{\partial}{\partial y}(DF) + \frac{\partial}{\partial z}(DG) \neq 0$ in B, then $y' = F$, $z' = G$ has no periodic orbits contained in B.*

Theorem 4. *There is no closed trajectory for system (11 – 12).*

Proof. We use the Dulac function to analyze whether equations (11) and (12) have a limit cycle. By making $D(y,z) = \frac{1}{y+z}$, then we get

$$\frac{\partial}{\partial y}(DF) + \frac{\partial}{\partial z}(DG) = -\frac{1 + G_1 + x + (G_4 + G_6)y + G_2 z}{(y+z)^2} < 0.$$

□

Hence, we do not have a limit cycle or closed trajectories and (y^*, z^*) is global. Because this is true, w's value is also global because equation (13) is a one dimensional system.

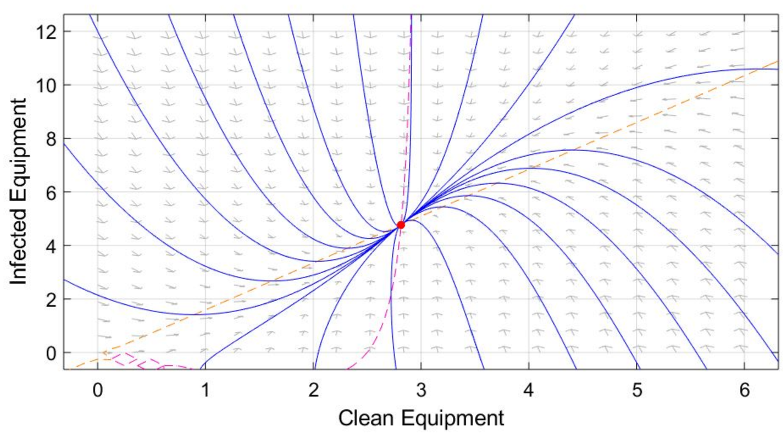

Figure 3: Global phase portrait of system (11 – 12) when $x^* = 0$

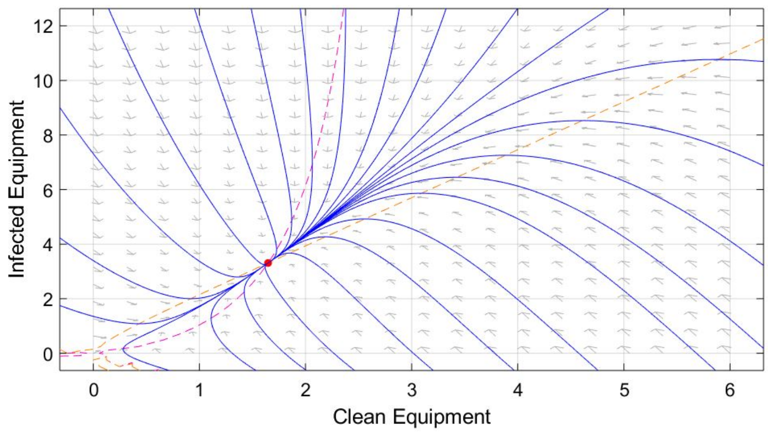

Figure 4: Global Phase Portrait of System (11 – 12) when $x^* = \frac{1}{A}(1 - G_0)$

Correspondingly, the phase portraits in Figure 3 and 4 show that the equilibria for equations (11 – 12) are globally stable for contamination-free equilibrium of soil and endemic equilibrium of soil respectively. This means that the trajectories will always head toward a positive equilibria for both contamination-free soil ($x^* = 0$) and contaminated soil equilibria $\left(x^* = \frac{1}{A}(1 - G_0)\right)$ regardless of the initial values.

5 Results

5.1 Impact of Control Parameters on R_0^S

We look at how the control parameters impact R_0^S. Figure 5 shows how R_0^S is influenced by t_A and ρ. The graph of R_0^S intersects with the flat plane where $R_0^S = 1$. The line that is generated by the intersection of planes is shown in Figure 6, which suggests that as the treatment rate of soil increases, the proportion of contaminated manure required to maintain $R_0^S = 1$ increases. We interpreted that as the proportion of infected manure increases, the need for treatment increases in order to get $R_0^S < 1$. The region above the line represent when the treatment is sufficient enough relative to the proportion of infected manure for R_0^S to be less than 1. The region below the line represents points for when the amount of treatment is not sufficient for the proportion of contamination in the manure. In this region, $R_0^S > 1$.

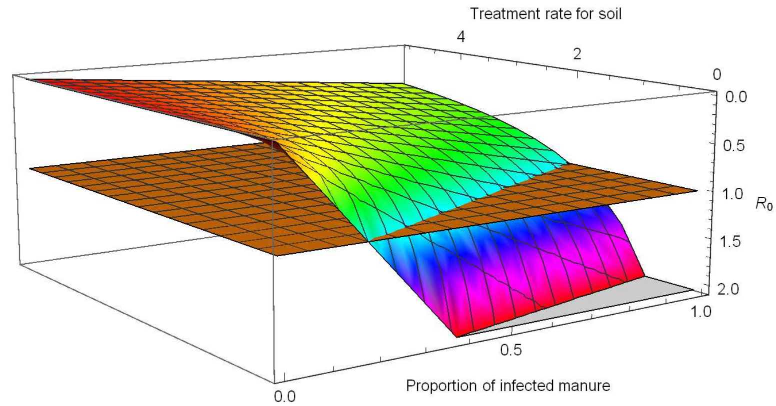

Figure 5: Graph of R_0^S as a function of t_A and ρ.

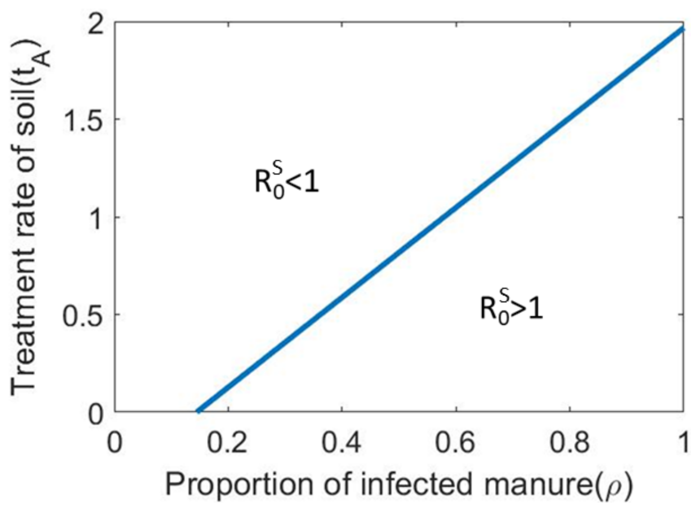

Figure 6: Bifurcation curve of R_0^S in $(\rho,\, t_A)$ – plane.

5.2 Sensitivity Analysis

We carry out a sensitivity analysis on R_0^S with respect to M, r, ρ, t_A, and d. In Figure 7 the most significant parameters are M, ρ, and r, which increase R_0^S. However, an increase in t_A and d will decrease the R_0^S, which is favored to reach a contamination-free equilibrium of soil.

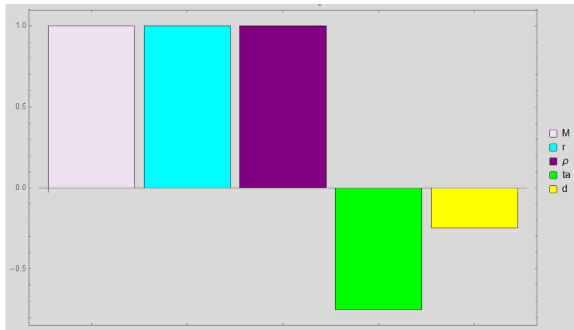

Figure 7: Sensitivity indices of R_0^S with respect to M, r, ρ, t_A, and d.

Figure 8 gives the sensitivity indices for the endemic equilibrium for P^*. As the values for r and M increase, P^* increases. When the values for t_A and d increase the value of P^* decreases. A smaller value for P^* means there is a smaller proportion of infected soil.

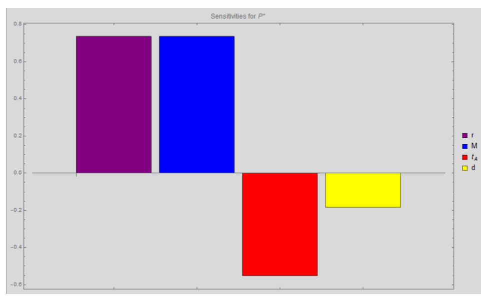

Figure 8: Sensitivity index of P^* with respect to r, M, t_A, and d.

Figure 9 describes the endemic equilibrium for E_c^* and E_I^*. The most significant parameter that increases equipment is δ_E, while r decreases equipment.

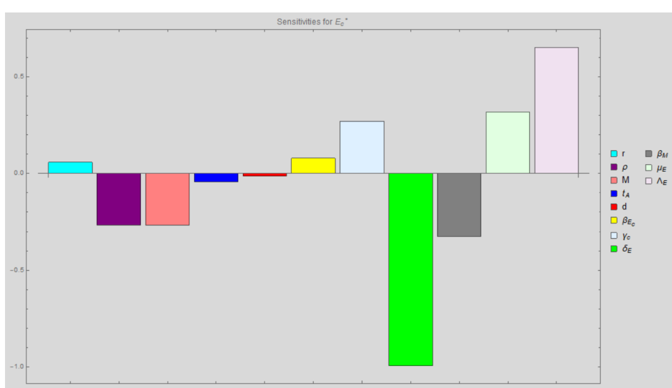

Figure 9: Sensitivity index of E_c^* with respect to r, ρ, t_A, d, β_{E_c}, γ_c, δ_E, β_M, μ_E, Λ_E

The sensitivity indices for contaminated lettuce, L_I^*, are very small when looking at the y–axis. The values for M and α are much larger than all the other parameters. Therefore, we separated M and α from the rest of them. We see that an increase in M, α has the greatest impact on L_I^* when compared to the other terms as can be seen in 10.

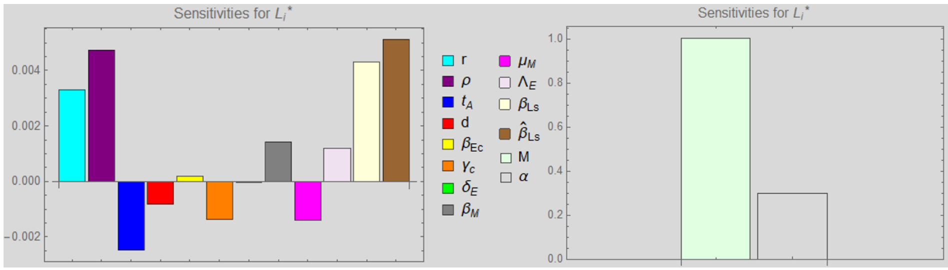

Figure 10: Sensitivity index of L_I^*. The values for parameters M and α are much larger than the rest so the two were extracted to their own plot (right), while the rest are shown on the left.

5.3 Simulations

We determine that increasing the cleaning rate of equipment will result in a decrease in the proportion of infected lettuce, as depicted in the nonlinear curve in Figure 11.

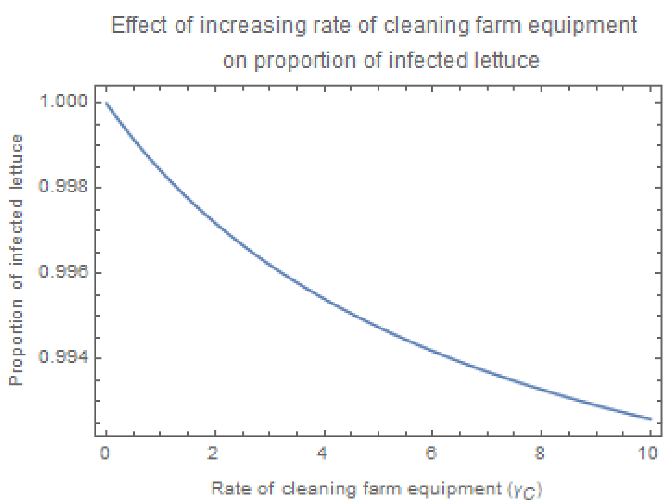

Figure 11: As the cleaning rate (γ_c) increases, the number of contaminated lettuce decreases.

Similarly, in Figure 12 we can show that the higher the rate of testing and treatment of soil for *E. coli*, the lower the proportion of infected equipment. Figure 13 shows this can lead to a higher population of healthy lettuce. Next, we run simulations of the proportion of infected soil, equipment, and lettuce over a time span of 20 days changing parameter values.

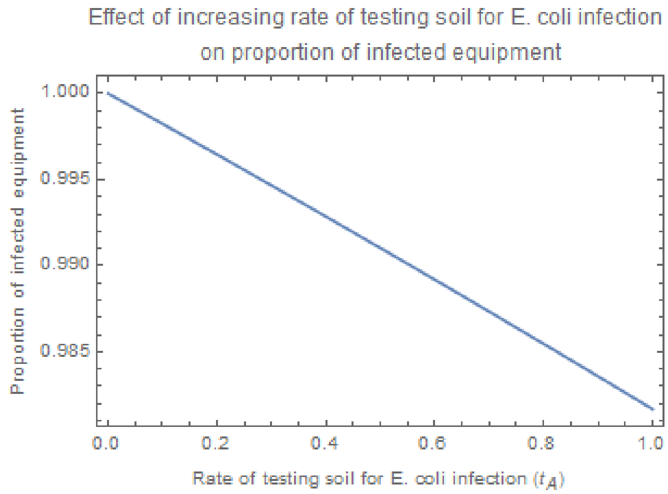

Figure 12: As treatment rate (t_A) increases, the number of contaminated equipment decreases.

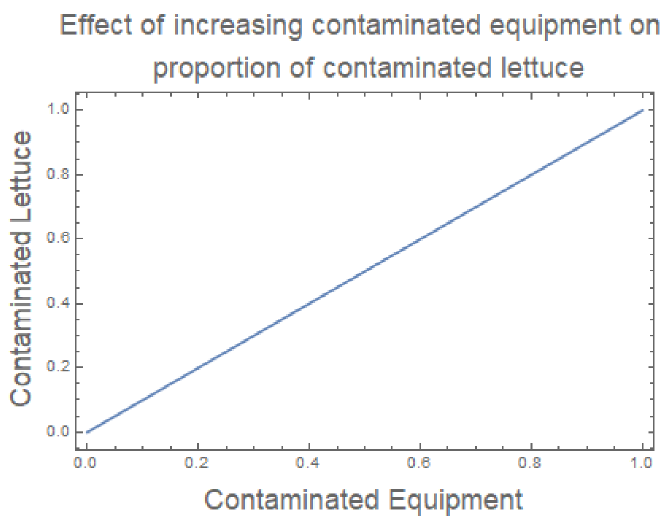

Figure 13: As contaminated equipment increases, contaminated lettuce increases.

In order to increase the amount of clean equipment, the removal rate of contaminated equipment due to inability to clean (δ_E) is increased. In Figures (14 – 16) the initial conditions for the proportion of infected soil, cleaned equipment, contaminated equipment, healthy lettuce and contaminated lettuce are $P_0 = 0.9$, $E_{c_0} = 9$, $E_{I_0} = 1$, $L_{S_0} = 1$, $L_{I_0} = 0$, respectively for all the simulations. In addition, the following parameters are $\rho = 9.91 \times 10^{-7}$, $r = 0.09$, $t_A = 0.9$, $\gamma_c = 10/7$, $\beta_{E_c} = 1$, $\beta_M = 0.5$, $\beta_E = 1$, $\mu_E = 1$, $\beta_{L_s} = 0.5$, $\mu_L = 0.9$, $\hat{\beta}_{L_S} = 1$, $\mu_E = 1$, $\Lambda_E = 10$, $\delta_E = 0.005$, and all other values (M, d, β_{L_S}) come from table 1. Hence $R_0^S < 1$ and as a result the proportion of infected soil eventually reaches zero as seen in Figure 14. In Figure 15, there are more clean equipment than contaminated equipment and Figure 16 shows that healthy lettuce is dominating.

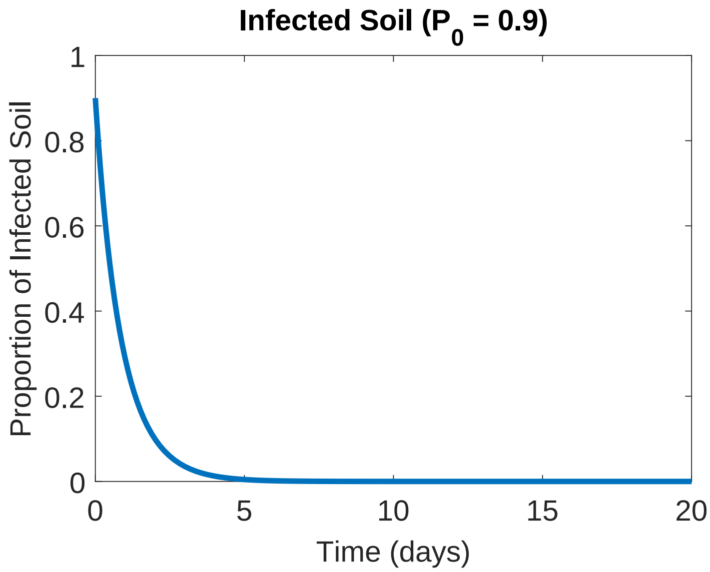

Figure 14: Infected soil over time.

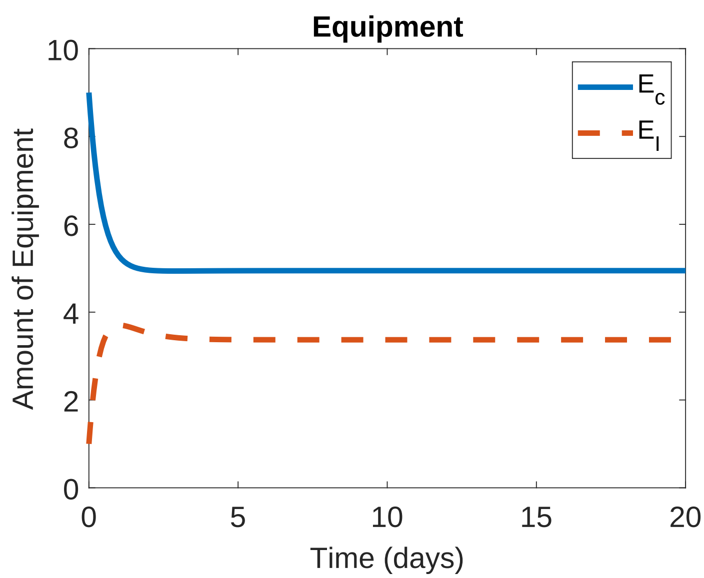

Figure 15: Equipment over time.

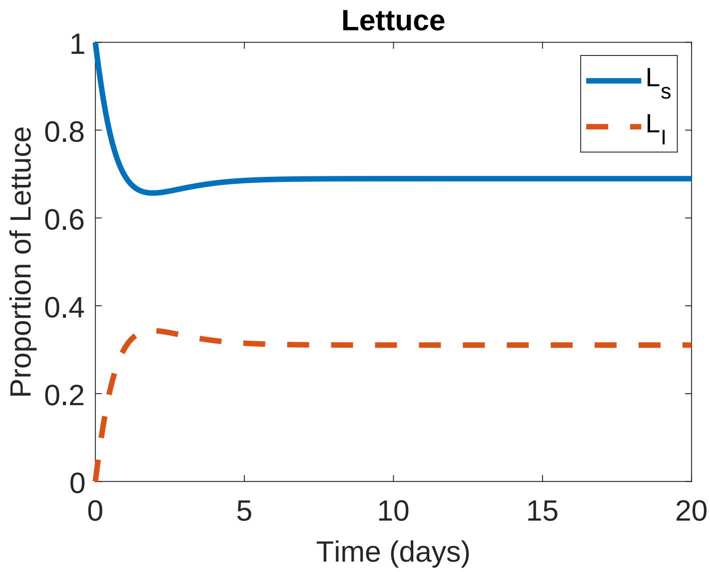

Figure 16: Lettuce over time.

6 Discussion

Similar studies investigated the probability of *E. coli* infection in lettuce, as well as prevention strategies, to reduce the chance of contaminated produce. Specifically, an article by Franz et al. (2008), focuses on modelling the likelihood that manure-amended soil from cattle infects lettuce. As a result, the "density of *E. coli* O157:H7 in manure-amended soil at the time of planting lettuce was most highly correlated to the storage time of the manure...and the initial concentration in manure..." [17]. This conclusion corresponds with our model in that once the manure becomes contaminated, then inevitably, the lettuce will become contaminated as well. Therefore, prevention efforts are suggested to take place early in the growing process. Examples would include proper composting and setting a minimum manure storage time to decrease the probability of infection. Another study concentrates on different models that help indicate the existence of *E. coli* O157:H7 in lettuce fields. When a scenario analysis was applied to a stochastic model, results showed that as time passed, the percentage of *E. coli* contaminated units decreased in the population [24]. In the same way, our simulation in Figure 14 shows that when $R_0^S < 1$, the proportion of infected soil decreases over time and eventually tends to 0.

Using our basic reproductive number for soil, $R_0^S = \frac{Mr\rho}{t_A+d}$, we can analyze the impact of testing and treatment of soil on the *E. coli* infection. Figures 5 and 6 indicate the t_A values for when the *E. coli* infection will die out in manure, meaning that more testing is needed to keep *E. coli* from becoming endemic. The positive correlation between ρ and t_A in Figure 6 demonstrates how much testing is required to reduce ρ in order for $R_0^S < 1$. In addition, Figure 7 illustrates that R_0^S is sensitive to t_A in a negative way such that when t_A is increased, the value of R_0^S becomes smaller. This is further evidence that testing and treatment of soil decreases *E. coli*; therefore, also impacting the yield of healthy lettuce. Furthermore, in analyzing Figure 8, we know that the proportion of infected soil, P, is influenced by t_A, similar to the impact of t_A on R_0^S. Not only does t_A affect R_0^S and P, but also the proportion of clean equipment. As previously demonstrated, when t_A is increased, P decreases which reduces the proportion of contaminated equipment, as shown in Figure 12. The collective influence of testing and treatment of contaminated soil proves that t_A is an important parameter in controlling the spread of *E. coli* and preventing the infection from reaching lettuce.

The cleaning of contaminated equipment is another parameter that can aid in increasing the yield of healthy lettuce. Based on Figures 9 – 10, we know that E_c is positively impacted by γ_c and L_I is negatively effected by γ_c which is what we would expect to see. This indicates that by increasing the cleaning rate of equipment, we can control the spread of the *E. coli* infection from reaching the population of healthy lettuce. As the cleaning rate of farm equipment increases, there is also a decline in the proportion of contaminated lettuce; therefore, increasing γ_c helps to prevent *E. coli* from transmitting across the lettuce field and promotes the growth of healthy lettuce, as shown in Figure 11.

In addition to t_A and γ_c being control parameters in our system, ρ is also a parameter that has an overall effect. If we can reduce the value of ρ, then the proportion of infected soil, number of contaminated equipment, and number of contaminated lettuce are all reduced.

There are limitations in our study due to the lack of information; therefore, it was challenging to accurately determine parameters and fully represent the dynamics of *E. coli* in soil. Although we were unable to find exact transmission parameters, we used our best judgment to estimate values based on information from literature. Additionally, it was difficult to find current standards for testing and treatment of the *E. coli* infection in soil. However, this information may be available in the future because the FDA announced a projected start date for farms to incorporate regular water and soil testing in spring 2019 [12]. Finally, to more accurately capture the complexity of the transmission of *E. coli* to lettuce, other factors could be considered in our model such as the role of cattle and irrigation as well as soil nutrition. Despite these limitations, the development of our model has still given us insight into how *E. coli* interacts with different elements of a farm setting including manure, soil, and equipment. The lack of studies on this topic indicates that in the future there needs to be more research on the role of equipment in the transmission of *E. coli* on farms.

In future studies, we would like to incorporate irrigation water to see how this mechanism contributes to the spread of *E. coli* in lettuce. The irrigation system is of high interest because of the water contamination risk. It is common for canals and rivers to be near farms, so it serves as a water source in the planting season. Sprinklers are used routinely during preharvest to assist the growth of lettuce and could potentially include serious contaminants like the Shiga toxin-producing *E. coli.* In 1995, the foodborne outbreak in lettuce was linked to *E. coli* bacteria found in irrigation water in Montana [13]. Cattle feces, surface runoff, and groundwater can enter into nearby water sources and further infect water that is used for irrigation [7]. Additionally, we can include more terms such as the infection in soil because of contaminated equipment into equation (1) to aid in capturing the complexity of *E. coli.* Finally, we could conduct a cost analysis on several scenarios within an *E. coli* outbreak. One scenario is what could happen to a farm when it receives bad publicity from an outbreak and the monetary impact of that on the farm. Another scenario we could analyze would be to optimize costs of testing and treatment of contaminated soil and cleaning of contaminated equipment for farmers in order for them to prevent an *E. coli* infection on the farm.

7 Acknowledgments

We would like to thank Dr. Carlos Castillo-Chavez, Founding and Co-Director of the Mathematical and Theoretical Biology Institute (MTBI), for giving us the opportunity to participate in this research program. We would also like to thank Co-Director Dr. Anuj Mubayi as well as Coordinator Ms. Rebecca Perlin and Management Intern Ms. Sabrina Avila for their efforts in planning and executing the day to day activities of MTBI. We also want to give special thanks to Bechir Amdouni, Dr. Baltazar Espinoza, Dr. Anarina Murillo, and Josean Valesquez-Molina. This research was conducted as part of 2018 MTBI at the Simon A. Levin Mathematical, Computational and Modeling Sciences Center (MCMSC) at Arizona State University (ASU). This project has been partially supported by grants from the National Science Foundation (NSF Grant MPS-DMS-1263374 and NSF Grant DMS-1757968), the National Security Agency (NSA Grant H98230-J8-1-0005), the Alfred P. Sloan Foundation, the Office of the President of ASU, and the Office of the Provost of ASU.

References

[1] Ayscue P., Lanzas C., Ivanek R., and Gröhn Y.T. Modeling On-Farm *Escherichia coli* O157:H7 Population Dynamics. *Foodborne Pathogens and Disease*, (4(6), pp.461-470), 2009.

[2] Bauer F., and Castillo-Chavez C. *Mathematical Models in Population Biology and Epidemiology*. Springer Science+Business Media, LLC, New York, NY, 2 edition, 2010.

[3] Arizona Farm Bureau. Yuma county: Americas winter vegetable capital. https://arizonaexperience.org/land/yuma-county-americas-winter-vegetable-capital.

[4] Carlos Castillo-Chavez, Horst Thieme, et al. Asymptotically autonomous epidemic models. 1994.

[5] CDC. Surveillance System Overview: National Shiga toxin-producing *Escherichia coli* (STEC) Surveillance Overview. *Atlanta, Georgia: US Department of Health and Human Services, CDC*, 2012.

[6] CDC. *E.coli (Escherichia coli)*. https://www.cdc.gov/ecoli/reporting-timeline.html, Dec 2014. Accessed: 2018-07-09.

[7] CDC. Irrigation Water Issues Potentially Related to Multistate E. coli Outbreak on Spinach. https://www.cdc.gov/nceh/ehs/activities/irrigation-water-ecoli-spinach.html, Oct 2014. Accessed: 2018-07-19.

[8] CDC. *E.coli (Escherichia coli)*. https://www.cdc.gov/ecoli/general/index.html, Feb 2018. Accessed: 2018-07-09.

[9] CDC. Symptoms and Sources of Food Poisoning. https://www.cdc.gov/foodsafety/symptoms.html-symptoms, Apr 2018. Accessed: 2018-07-09.

[10] Fairbrother J.M. and Nadeau E. *Escherichia coli*: On-Farm Contamination of Animals. *Rev Sci Tech*, (25(2), pp.555-559), 2006.

[11] FDA. Final Qualitative Assessment of Risk to Public Health from On-Farm Contamination of Produce. https://www.fda.gov/downloads/Food/FoodScienceResearch/RiskSafetyAssessment/UCM470780.pdf, Nov 2015. Accessed: 2018-07-25.

[12] FDA. FSMA Final Rule on Produce Safety. https://www.fda.gov/food/guidanceregulation/fsma/ucm334114.htm, Jun 2018. Accessed: 2018-07-25.

[13] Fonseca, J.M. and Fallon, S.D. and Sanchez, CA and Nolte, KD. Escherichia coli survival in lettuce fields following its introduction through different irrigation systems. *Journal of Applied Microbiology*, 110(4):893–902, 2011.

[14] Food and Drug Administration. Commodity Specific Food Safety Guidelines for the Lettuce and Leafy Greens Supply Chain. https://www.fda.gov/downloads/food/guidanceregulation/ucm169008.pdf, Apr 2006. Accessed: 2018-07-05.

[15] Food and Drug Administration. Food Safety Modernization Act Final Rule on Produce Safety. https://www.fda.gov/Food/GuidanceRegulation/FSMA/ucm334114.htm, July 2018. Accessed: 2018-07-14.

[16] Food Chain Crisis Management Framework. Preventing Escherichia coli in Food. http://www.fao.org/fileadmin/user_upload/fcc/news/1_FAO_Preventing-E.Coli-inFood_FCC_2011.06.23.pdf. Accessed: 2018-07-09.

[17] Franz E., Semenov A.V., Van Bruggen A.H.C. Modelling the contamination of lettuce with *Escherichia coli O157: H7* from manure-amended soil and the effect of intervention strategies. *Journal of applied microbiology*, 5(105, pp.1569-1584), 2008.

[18] Interagency Food Safety Analytics Collaboration. Foodborne illness source attribution estimates for 2013 for Salmonella, *Escherichia coli O157*, Listeria monocytogenes, and Campylobacter using multi-year outbreak surveillance data, United States. Dec 2017.

[19] Iowa State University Extension and Outreach. Replacement Strategies for Farm Machinery. https://www.extension.iastate.edu/agdm/crops/html/a3-30.html, May 2018. Accessed: 2018-07-15.

[20] Kansas Healthy Yards and Communities. Home Soil Testing for *E. Coli*, note=Accessed: 2018-07-10, howpublished="http://www.kansashealthyyards.org/index.php?option=com_allvideoshare&view=video&slg=home-soil-testing-for-e-coli&Itemid=340",.

[21] Kerns D.L., Matheron M.E., Palumbo J.C., Sanchez C.A., Still D.W., Tickes B.R., Umeda K., Wilcox M.A. Guidelines for Head Lettuce Production in Arizona. https://cals.arizona.edu/crop/vegetables/cropmgt/az1099.html, Feb 1999. Accessed: 2018-07-09.

[22] Levins R. Some Demographic and Genetic Consequences of Environmental Heterogeneity for Biological Control. *The ISME journal*, (pp. 173-183), Feb 2011.

[23] Mayo Clinic Staff. *E. coli.* https://www.mayoclinic.org/diseases-conditions/e-coli/symptoms-causes/syc-20372058. Accessed: 2018-07-10.

[24] McKellar R.C., Peréz-Rodríguez F., Harris L.J., Moyne A-L., Blais B., Topp E., Bezanson G., Bach S., Delaquis P. Evaluation of different approaches for modeling Escherichia coli O157: H7 survival on field lettuce. *International journal of food microbiology*, 184:74–85, 2014.

[25] Painter J.A., Hoekstra R.M., Ayers T., Tauxe R.V., Braden R., Angulo F.J., Griffin P.M. Attribution of Foodborne Illnesses, Hospitalizations, and Deaths to Food Commodities by using Outbreak Data, United States, 19982008. *Emerging Infectious Diseases*, (3(19)), Mar 2013.

[26] Solomon E.B., Yaron S., Matthews K.R. Transmission of Escherichia coli O157:H7 from Contaminated Manure and Irrigation Water to Lettuce Plant Tissue and Its Subsequent Internalization. *Applied and Environmental Microbiology*, (1(68), pp. 397-400), Jan 2002.

[27] Strohbehn C., Mendonca A., Wilson L., Domoto P., Smith M. On-farm Food Safety: Cleaning and Sanitizing Guide. *Human Sciences Extension and Outreach Publications*, (pp. 102), 2013.

[28] Washington State University Extension. 2011 Cost Estimates of Producing Fresh Market Field-Grown Head Lettuce in Western Washington. `http://cru.cahe.wsu.edu/CEPublications/FS081E/FS081E.pdf`, Nov 2012. Accessed: 2018-07-14.