AJUR

American Journal of
Undergraduate Research

Volume 18 | Issue 3 | December 2021

www.ajuronline.org

Print Edition ISSN 1536-4585
Online Edition ISSN 2375-8732

AJUR

American Journal of
Undergraduate Research

Volume 18 | Issue 3 | December 2021 *www.ajuronline.org*

American Journal of Undergraduate Research (AJUR) *is a national, independent, peer-reviewed, open-source, quarterly, multidisciplinary student research journal. Each manuscript of AJUR receives a DOI number. AJUR is archived by the US Library of Congress. AJUR was established in 2002, incorporated as a cheritable not-for-profit organization in 2018. AJUR is indexed internationally by EBSCO and Crossref with ISSNs of 1536-4585 (print) and 2375-8732 (web).*

An Optimal Control Experiment for an SEIRS Epidemiological Model

Tanner Snyder, Ryan Nierman*

Department of Mathematics and Computer Science, Madonna University, Livonia, MI

https://doi.org/10.33697/ajur.2021.047

*Student: tsnyder213@my.madonna.edu**
Mentor: rnierman@madonna.edu

ABSTRACT
This work studies an optimal control model for a discrete-time Susceptible/Exposed/Infective/Removed/Susceptible (SEIRS) deterministic epidemiological model with a finite time horizon and changing population. The model presented converts a continuous SEIRS model that would typically be solved using differential equations into a discrete model that can be solved using dynamic programming. The discrete approach more closely resembles real life situations, as the number of individuals in a population, the rate of vaccination to be applied, and the time steps are all discrete values. The model utilizes a previously developed algorithm and applies it to the presented SEIRS model. To demonstrate the applicability of the algorithm a series of numerical results are presented for various parameter values.

KEYWORDS
Control; Cost; Discrete; Disease; Epidemiology; Minimization; Modeling; Optimality; SEIRS; Vaccination

INTRODUCTION
Epidemiological modeling relies on the usage of numerous epidemiological classes. The most common are described by M, S, E, I, and R. M represents the class of individuals who have passive immunity, S represents the susceptible class (individuals who are capable of becoming infected), E represents the exposed class (individuals who are infected, but not yet infectious), I represents the infected class (those who are infectious and capable of spreading the disease to a member of the susceptible class), and R represents the recovered or removed class. This class may consist of those with either permanent infectious-acquired immunity, temporary immunity, or those who have died/emigrated, and therefore, have been removed from the population. Some of the most common epidemiological models include: MSEIR, MSEIRS, SEIR, SEIRS, SIR, SIRS, SEI, SEIS, SI, and SIS.[2, 4, 8, 15, 16] **Figure 1** shows the general transfer diagram for a MSEIR model.[10]

The research presented explores a Susceptible/Exposed/Infective/Removed/Susceptible (SEIRS) model. Typically, epidemiological models are represented by one or more ordinary or partial differential equations, which are then used to model the spread of the epidemic or solve questions of optimality.[9, 17] In the case of the research presented, the continuous model is converted to a discrete model in which a nonlinear programming method for finding the optimal control sequence is applied for the discrete-time deterministic SEIRS epidemic. The algorithm will seek to find the minimum cost necessary to control an epidemic under a finite time horizon and discrete state space. The numerical results and corresponding graphs are generated using a python script.[A]

[A]The python script can be found on GitHub at https://github.com/tsnyder213/SEIRS-Optimality

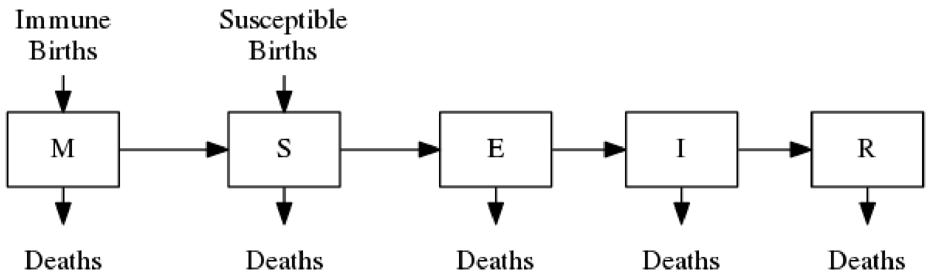


Figure 1. A General Transfer Diagram for the MSEIR Model.

AN SEIRS EPIDEMIC OPTIMAL CONTROL PROBLEM

A typical Susceptible/Exposed/Infective/Removed/Susceptible epidemiological model is presented using a system of differential equations. The total population N is divided into multiple subpopulations/classes, namely, S, E, I, and R, where S represents the susceptibles, E represents the exposed class (or those that are infected, but not yet infectious), I represents the infectives, and R represents those that have been removed from the population due to death as a result of the disease. A basic SEIRS model assumes that the total population, N, does not allow for natural births/deaths nor immigration/emigration, therefore individuals are only able to move between subpopulations. Having said that, the model utilized in the research explores a more advanced model that allows for a changing population. Additionally, the model assumes that it is possible for an individual to become infectious, recover from the disease, and reenter the susceptible class, which is why the model presented represents an $SEIRS$ model, rather than an $SEIR$ one.

The following systems of differential equations represents a common SEIRS model without natural births/deaths or immigration/emigration included.[1, 7, 13]

$$\frac{dS}{dt} = -\frac{\beta IS}{N} + \tau I \qquad \textbf{Equation 1.}$$

$$\frac{dE}{dt} = \frac{\beta IS}{N} - \mu E \qquad \textbf{Equation 2.}$$

$$\frac{dI}{dt} = \mu E - \gamma I - \tau I \qquad \textbf{Equation 3.}$$

$$\frac{dR}{dt} = \gamma I. \qquad \textbf{Equation 4.}$$

In this model, $\beta \in [0, 1]$ represents the probability of infection when an infective comes into contact with a susceptible, $\tau \in [0, 1]$ represents the rate of recovery, $\mu \in [0, 1]$ represents the rate of people leaving the exposed class and entering the class of infectives, and $\gamma \in [0, 1]$ represents the rate of those removed from I as a result of dying from the disease.[5, 8, 9, 12, 14–16] The term

$$\frac{\beta IS}{N} \qquad \textbf{Equation 5.}$$

is referred to as the interaction term because it represents the interaction between an infective and a susceptible and the effect that β has on transmission during that interaction.[1, 10, 13, 15, 16]

To allow the total population N to change due to natural births/deaths and/or immigration/emigration, the previous system of differential equations outlined in **Equations 1-4** becomes:

$$\frac{dS}{dt} = \phi N - \frac{\beta IS}{N} + \tau I - \delta_1 S \qquad \textbf{Equation 6.}$$

$$\frac{dE}{dt} = \frac{\beta IS}{N} - \mu E - \delta_2 E \qquad \textbf{Equation 7.}$$

$$\frac{dI}{dt} = \mu E - \gamma I - \tau I - \delta_3 I \quad \text{Equation 8.}$$

$$\frac{dR}{dt} = \gamma I + \delta_1 S + \delta_2 E + \delta_3 I, \quad \text{Equation 9.}$$

where ϕ a rate of population growth due to natural births or immigration. This term can be referred to as the recruitment rate of the susceptibles. It is assumed that all natural births and immigration result in new susceptibles, therefore, no one can enter the population and immediately be entered into the class of exposed or infectious individuals. Furthermore, δ_1, δ_2, and δ_3 represent the rates of natural deaths or emigration unrelated to the disease. It is assumed that $\phi \in [0,1]$, $\delta_1 \in [0,1], \delta_2 \in [0,1]$, and $\delta_3 \in [0,1]$.[7]

Having established the continuous representation for an SEIRS model, it is now necessary to convert the model into a discrete model, which can be used to establish an optimal treatment strategy to minimize the cost of controlling the spread of the epidemic. By relying on the use of discrete mathematics, the model will be able to overcome issues with convergence and ill-behaved functions. Discrete modeling is more realistic than continuous models since it relies on discrete time steps and individuals within a population.[15, 16] Therefore, instead of using ordinary differential equations to express changes in epidemiological classes over time (i.e., the rates at which individuals are moving between subpopulations), S, E, I and R are going to be based solely on previous discrete time steps. The control setting is an epidemic process with a finite time horizon and discrete state space. Let the epidemic evolve over integral decision times $\{T \in \mathbb{N} \mid 0 \leq T < \infty\}$. At the start of each time interval $(t, t+1)$, a decision, u_t, specifying the rate of vaccination (control) to apply must be made. Hence, u_t is an element of a vector u of decisions of length T. The number of possible control rates is based on the number of partitions/partial rates of vaccination (control) allowed by the model.[10, 15, 16]

Allowing for a changing population, and converting from continuous to discrete mathematics, creates the following model:

$$S_{t+1} = S_t + \lfloor \phi(S_t + I_t + E_t) \rceil - \left\lfloor \frac{\beta I_t S_t}{S_t + I_t + E_t} \right\rceil + \lfloor \tau I_t \rceil - \lfloor \delta_1 S_t \rceil \quad \text{Equation 10.}$$

$$E_{t+1} = E_t + \left\lfloor \frac{\beta I_t S_t}{I_t + S_t + E_t} \right\rceil - \lfloor \mu E_t \rceil - \lfloor \delta_2 E_t \rceil \quad \text{Equation 11.}$$

$$I_{t+1} = I_t + \lfloor \mu E_t \rceil - \lfloor \gamma I_t \rceil - \lfloor \tau I_t \rceil - \lfloor \delta_3 I_t \rceil \quad \text{Equation 12.}$$

$$R_{t+1} = \lfloor \gamma I_t \rceil + \lfloor \delta_1 S_t \rceil + \lfloor \delta_2 E_t \rceil + \lfloor \delta_3 I_t \rceil \quad \text{Equation 13.}$$

The notation, $\lfloor \cdot \rceil$, represents a rounding function (formed from half of the ceiling function notation and half of the floor function notation) that forces the result to be a discrete value.[15]

The presented discrete model relies on a couple of core assumptions. First, the model relies on homogeneous mixing in which exactly one infective comes into contact with exactly one susceptible during each time step $(t, t+1)$. Therefore, it is only possible to spread the disease if the individual is actually infectious and not just infected (i.e., an individual in the exposed class cannot infect a susceptible). Second, the model assumes that the time it takes for an individual to transition from one subpopulation to another is negligible. Third, natural births and immigration results in only susceptible individuals, but individuals are able to emigrate or die from natural causes from any class. Therefore, R represents individuals who have been removed from the population due emigration, natural deaths, or deaths as a result of the epidemic. Furthermore, at time $t = 0$, $R_0 = 0$, but at time $t > 0$, R_t does not influence S_{t+1}, I_{t+1}, or E_{t+1}. Therefore, R_t does not effect the interaction term given by:

$$\left\lfloor \frac{\beta I_t S_t}{S_t + I_t + E_t} \right\rceil. \quad \text{Equation 14.}$$

Therefore the denominator for the interaction term is presented as $I_t + S_t + E_t$, rather than N.[3, 6, 10, 15, 16]

In order to control the spread of the epidemic, one needs to impose a rate of vaccination to lower the transmission parameter, B. Letting u represent the control vector and u_t represent the rate of vaccination used during each time step $(t, t+1)$, B can be defined as $B = \beta(u)$. Therefore, B is a known decreasing function in u. For binary control,

$$\beta(u) = \begin{cases} \beta(0) = B_{max} \\ \beta(\sigma) = B_{min}, \end{cases} \tag{Equation 15.}$$

where $B_{max} > B_{min}$. Since B_{max} and B_{min} are probabilities of transmission, both $B_{max}, B_{min} \in [0, 1]$. For the purposes of the binary model presented, $u_t \in \{0, \sigma\}$, where one can either administer no vaccine (a zero vaccination rate) or administer the maximum vaccination rate, σ. Therefore, a relationship exists between whether a control is administered and the probability, β, of infection. If the rate of vaccination is zero, then the probability that the infection spreads should be at its highest, B_{max}. If the rate of vaccination administered is at its maximum, then the probability that the infection spreads should be at its lowest, B_{min}. The latter increases to reflect that the cost of eradicating a disease increases as its prevalence decreases, since it is harder to control a small and decreasing fraction of infectives.[3]

There is a cost associated with administering the vaccine in order to control the current infectives, but there is also a cost associated with the cost per new infective. In the case of continuous mathematics, this can be modeled by minimizing the objective (loss) function

$$L = \int_0^T \left[c_1 s + c_2 \gamma\right] e^{-\alpha t} dt, \tag{Equation 16.}$$

where c_1 is the social cost per infective, c_2 is the cost of control per unit level of program effort, α is a discount rate, and T represents a finite time horizon.[18]

Since the approach presented in this paper relies on discrete mathematics, the loss function used to calculate the cost at a particular time step becomes:

$$L_t(u) = \left\lfloor \frac{\beta I_t S_t}{S_t + I_t + E_t} \right\rceil CI + u_t \frac{S_t + I_t + E_t}{I_t + 1}, \tag{Equation 17.}$$

where CI is the cost per new infective and u_t is the control vector value at the current time step. If the number of partitions associated with the possible rates of vaccination is changed, u_t will update accordingly to allow for the change in granularity; thereby, influencing the cost associated with controlling the epidemic.

The cumulative loss function used to calculate the cumulative cost, J_t, is defined as follows:

$$J_t = \begin{cases} 0 & t = 0 \\ J_{t-1} + L_t & t \geq 1. \end{cases} \tag{Equation 18.}$$

Therefore,

$$J_{t+1} = J_t + \left\lfloor \frac{\beta(u_t) I_t S_t}{S_t + I_t + E_t} \right\rceil CI + u_t \frac{S_t + I_t + E_t}{I_t + 1}. \tag{Equation 19.}$$

Therefore, the problem is to find the optimal control vector u_t for $\{T \in \mathbb{N} \mid 0 \leq T < \infty\}$ that minimizes J_t.[3, 15]

The problem can be expanded to allow for k-ary control in which the rate of vaccination does not need to be restricted to either no vaccination or full vaccination. Therefore $u_t = \left\{\frac{j}{k}\sigma, j = 0, ..., k\right\}$.[3] In order to apply partial rates of vaccination, two assumptions must be made: First, it is assumed that there is a linear association between the cost and the rate of vaccination being applied. Second, it is assumed that B remains a decreasing function of u, where B must be able to take on an infinite number of possible values rather than two points.[3]

Using the two intersection points, $(0, B_{max})$ and (σ, B_{min}), one can find the linear representation for β, as shown in **Equation 20.**[15, 16]

$$\beta(u) = \left(\frac{B_{min} - B_{max}}{\sigma}\right)u + B_{max} \qquad \text{Equation 20.}$$

Since the algorithm that will be utilized to solve the minimization problem will rely on an "intelligent" search of the space of feasible solutions, a state space needs to be established. Let $T = (V, A)$ be a graph with the following node and arc set: $V : (I_t, J_t)$ and $A : ((I_t, J_t), (I_{t+1}, J_{t+1}))$, such that:

$$I_{t+1} = I_t(u), u \in \{0, \sigma\} \qquad \text{Equation 21.}$$

$$J_{t+1} = J_t + L(I_t, u). \qquad \text{Equation 22.}$$

Figure 2 represents the initial time step in T for a binary control tree. The root is defined by I_0 and J_0, where I_0 represents the initial number of infectives and $J_0 = 0$ since the initial cumulative cost is zero. The left child, denoted by $u = 0$, represents a vaccination rate of zero. The right child, denoted by $u = \sigma$, represents applying the maximum rate of vaccination. The resulting nodes (I_1^0, J_1^0) and (I_1^σ, J_1^σ) represent the two children of the root, where the subscripts represent the current time step, $t = 1$, and the the superscripts, 0 or a σ, represents whether the vaccination rate was zero or maximum. After the initial time step, the notation is simplified. I_i is defined to be the number of infectives at node i and J_i represents the cumulative cost at node i (i.e., the sum of all of the costs at each node on the path from the root to node i), and t_i is the level of node i. This assumes that tree T grows downward.[15]

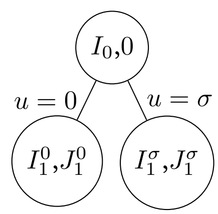


Figure 2. Initial Time Step of a Binary Control Tree.

To allow for additional rates of vaccination (partial levels of control), $u_t = \{\frac{j}{k}\sigma, j = 0, ..., k\}$. Therefore, the resulting decision tree is formed by a root with $k + 1$ children. Each child corresponds to a possible partial rate of vaccination, as shown in **Figure 3**. As in the case of the binary control decision tree, **Figure 3** represents the initial time step in T, where the root is defined by I_0 and $J_0 = 0$. The root's $k + 1$ children represents a zero rate of vaccination, maximum rate of vaccination, and $k - 1$ rates of partial control. As before, each child node has a subscript denoting the current time step, and a superscript corresponding to the amount of control being applied.[15]

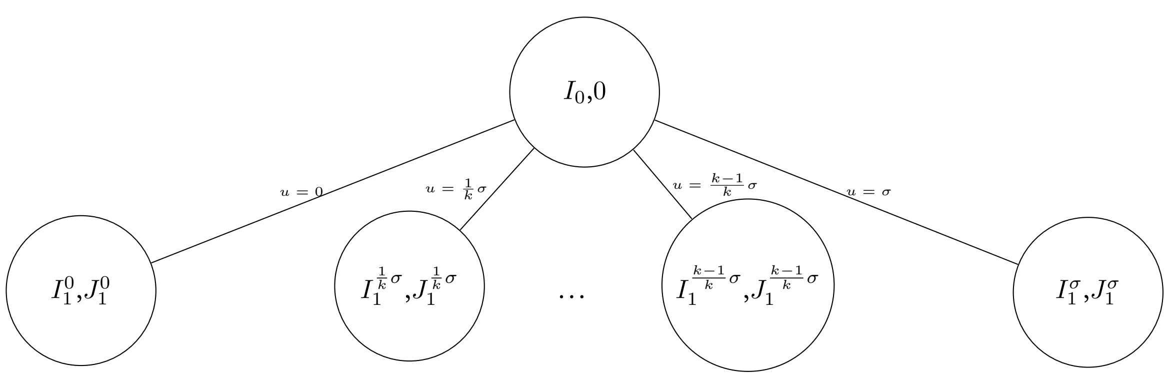


Figure 3. Initial Time Step of a K-ary Control Tree.

In order to avoid an exponential expansion of the search space, especially as the number of partial vaccination rates goes up, a method of pruning subtrees should be utilized to eliminate paths that lead to leaves that are guaranteed not to yield optimal solutions.

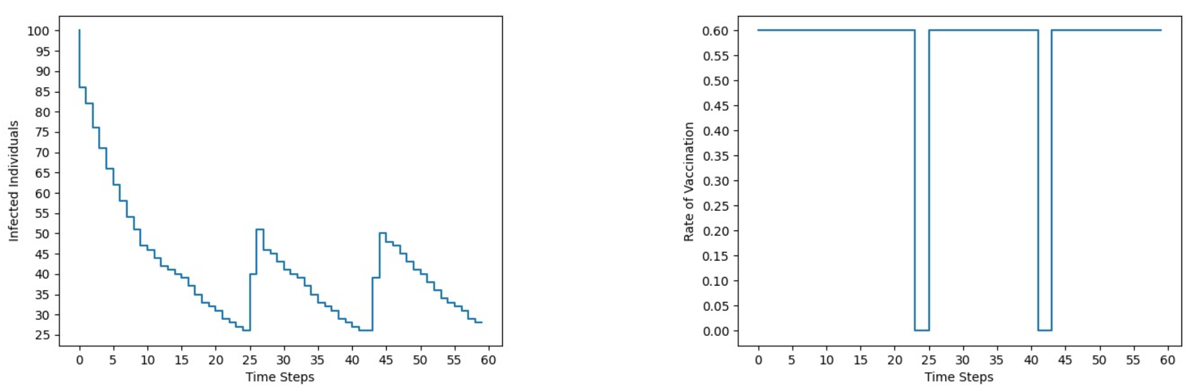
Let T be a state space tree. Let $i,j \in V(T)$ s.t. i is a node at level t_i and j is a node at level t_j, $t_i \geq t_j$, $I_i = I_j$, and $J_j \leq J_i$. Then T_i, the subtree generated by node i, can be pruned.[15]

The algorithm utilized to find the optimal solution seeks to minimize the cumulative cost of controlling an epidemic over a finite time horizon. The algorithm begins with an initial number of infectives I_0 and an initial cumulative cost $J_0 = 0$. The loss function is calculated for each of the possible levels of vaccination. A comparison is made at each step according to the pruning technique above, so that large portions of the tree can be pruned. Pruning subtree T_i results in a significant decrease in the number of paths that must be searched. Finally, the algorithm compares the resulting cumulative costs at the leaves of the remaining paths to find the optimal solution.

NUMERICAL RESULTS

To demonstrate the utility of the algorithm, a series of numerical results are presented. Unless otherwise specified, the parameter values are: $T = 60, N_0 = 500, S_0 = 395, I_0 = 100, E_0 = 5, u_t \in [0, 0.6], B_{max} = 0.75, B_{min} = 0.15, \tau = 0.13, \gamma = 0.05$, $\sigma = 0.6$, and $CI = 0.65$. For each scenario, two discrete graphs are presented. The graph on the left represents the number of infected individuals (infectives) at each time step. The rate of vaccination graph on the right represents the optimal control solution which minimizes the cost of controlling the epidemic over the finite time horizon for the provided parameters.

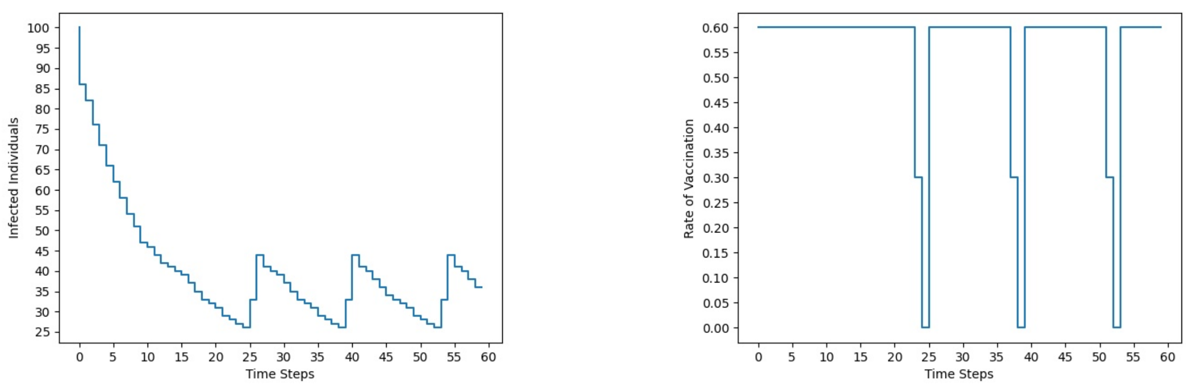
The algorithm is initially applied to a binary control system, in which the rate of vaccination must be either 0 or 0.6. **Figure 4** shows the change in infected individuals and the applied rate of vaccination for a binary control system. Based on the provided parameters, the optimal solution to minimize the cost of controlling the epidemic is to consistently vaccinate at the maximum rate of 0.6 for the entire time horizon except for times $t = 24, t = 25, t = 42$, and $t = 43$. The initial vaccination effort is able to get the infectives down from 100 to 28 after the first 23 time steps. Once the vaccination effort is stopped, the number of infectives begins to increase very quickly to 51 individuals by $t = 27$, therefore, requiring the vaccination efforts to resume. The same scenario occurs at $t = 44$.



(a) Number of Infectives **(b)** Rate of Vaccination

Figure 4. Binary Control Experiment.

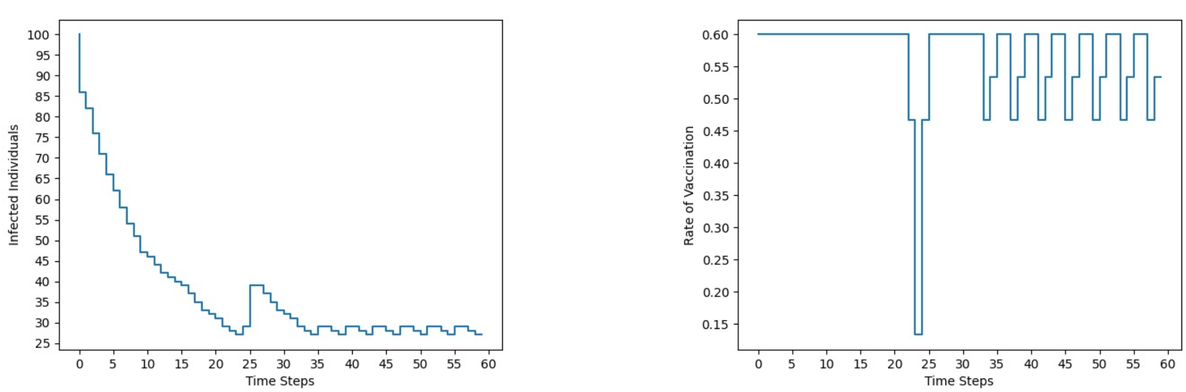
To demonstrate the algorithm's effectiveness for non-binary control scenarios, a series of k-ary control experiments are completed. **Figures 5-7** demonstrates the effect of increasing the granularity of the vaccination efforts to allow for partial rates of vaccination to be implemented. **Figure 5** demonstrates the effect of adding one additional rate of vaccination. Therefore, instead of being restricted to either applying no vaccination or the maximum rate of vaccination, the 3-ary control model also has the ability to choose to administer a $\frac{1}{2}$ vaccination rate. Since the maximum allowable rate of vaccination is 0.6, a $\frac{1}{2}$ vaccination rate would correspond to a rate of 0.3. Hence, $u_t = \{0, 0.3, 0.6\}$. The optimal control solution shown is similar to the binary control. The algorithm administers the maximum rate of vaccination until time $t = 23$ in order to rapidly decrease the number of infectives, after which an oscillatory pattern begins, toggling off and on the control as needed to maintain control of the epidemic, while minimizing the cost. At time $t = 24$, $t = 36$, and $t = 52$, a 0.3 vaccination rate is implemented; therefore, the algorithm is taking advantage of the increase in granularity available. As far as the number of infectives goes versus time, the increase in granularity decreases the heights of each of the peaks and results in three peaks, rather than two as in the binary experiment.



(a) Infectives **(b)** Rate of Vaccination

Figure 5. 3-Ary Control Experiment.

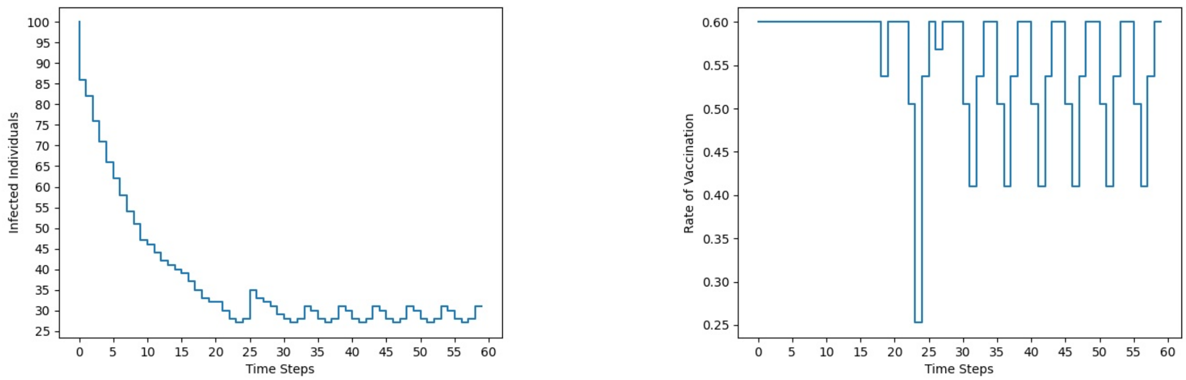
Increasing the granularity to 10 levels of partial vaccination results in the outputs shown in **Figure 6**. As one can observe, the graph of the number of infectives begins to smooth out. There is only one significant peak, which occurs at time $t = 26$ and some rippling between $t = 35$ and $t = 60$. The graph of the rate of vaccination demonstrates that the optimal solution is to control at the maximum rate for the first 22 time steps, before letting off to a vaccination rate of 0.47 at $t = 23$ and 0.13 at $t = 24$. Then, as the number of infectives increases, the need to apply a greater rate of vaccination occurs, therefore, the vaccination rate at $t = 25$, $t = 26$, and $t = 27$ are 0.13, 0.47, and 0.6, respectively. At this point, the rate of control applied oscilliates between 0.47 and 0.6 for the duration of the time horizon.



(a) Infectives **(b)** Rate of Vaccination

Figure 6. 10-Ary Control Experiment.

Finally, **Figure 7** represents 20 levels of control. As expected, the graph of the infectives continues to smooth out and the rippling further disappears. This time the largest peak only gets back up to 35 infectives at time $t = 26$. This is due to the increase in granularity as demonstrated in the rate of vaccination output. A similar oscillatory pattern occurs in the 20-ary scenario as it did in the 10-ary scenario once the initial control of the epidemic occurs. Therefore, the rate of vaccination is stepped up only as needed to regain control of the small, rippling peaks.

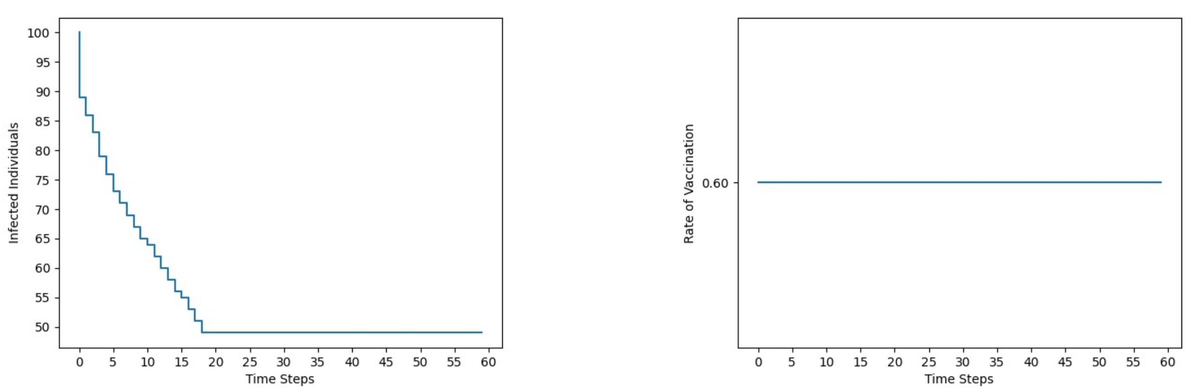


(a) Infectives **(b)** Rate of Vaccination

Figure 7. 20-Ary Control Experiment.

Having explored the effects of increasing the granularity of the vaccination efforts, an investigation into the effects of changing the other parameters is carried out for the binary control experiment.

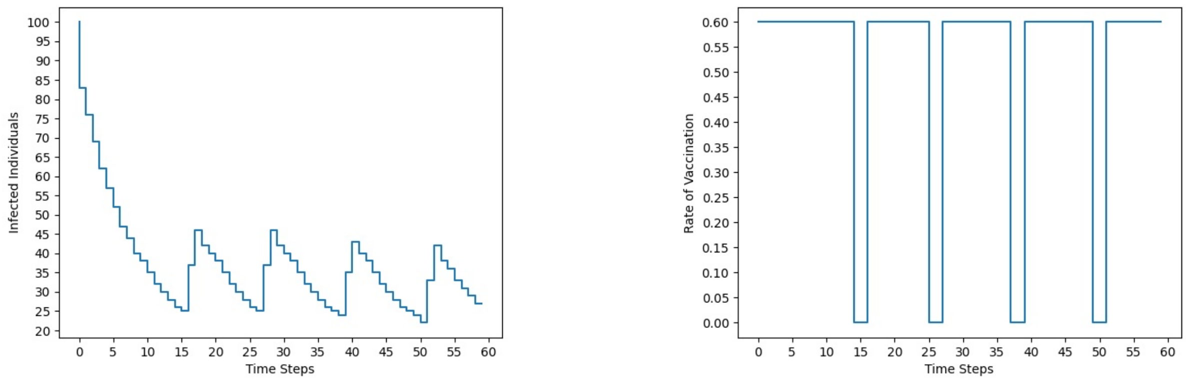
To begin, γ, the rate of death as a result of the infection is studied. Until this point, γ was set to 0.05. **Figure 8** shows the effect of decreasing γ to 0.02, thereby making the disease less deadly. In this case, it becomes difficult to control the epidemic, since less individuals are leaving I and entering R. Therefore, the number of infectives which can transmit the disease stays higher. Therefore, in order to get and maintain control of the spread of the epidemic, the optimal solution is to continuously vaccinate at the maximum rate. As one can observe from the graph of the infectives, it is necessary to control at the maximum rate to decrease the number of infectives from 100 to 49, where it plateaus for the duration of the time horizon ($t = 19$ through $t = 60$).



(a) Infectives **(b)** Rate of Vaccination

Figure 8. Binary Control Experiment with $\gamma = 0.02$.

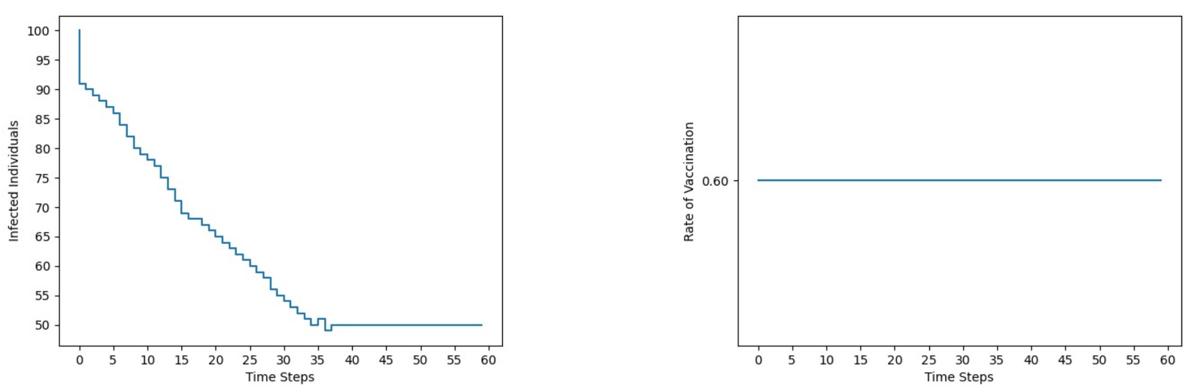
Figure 9 demonstrates the effect of increasing the deadliness of the disease to $\gamma = 0.08$. In this case, less vaccination is required to control the epidemic, since the disease is decreasing the number of individuals in I and moving them to R. Therefore, the disease's ability to kill its hosts acts as an internal control for the epidemic, so external intervention becomes less necessary that it did for $\gamma = 0.05$. A general oscillatory pattern occurs with $\gamma = 0.08$, but there are shorter periods in which the maximum rate of control is needed.



(a) Infectives **(b)** Rate of Vaccination

Figure 9. Binary Control Experiment with $\gamma = 0.08$.

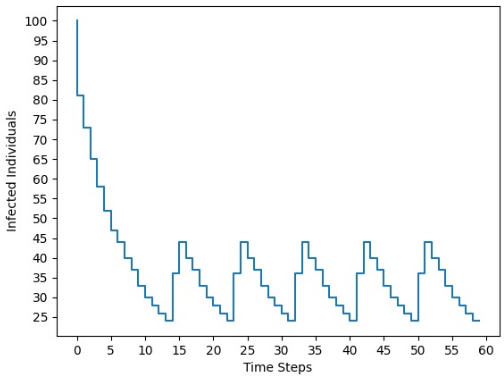
Next, the effect of the recovery rate is analyzed. Until now the recovery rate, τ, has been set to 0.13. In **Figures 10-11**, the effects of altering the recovery rate are shown. **Figure 10** represents the effect of decreasing τ from 0.13 to 0.08. In this case, it becomes harder to control the epidemic, because individuals in the infective class are no longer recovering as quickly and reentering the class of susceptibles. Therefore, there are more infectives available at each time step able to transmit the infection to a susceptible. Therefore, it becomes necessary to continuously vaccinate at the maximum rate to maintain control of the epidemic and then maintain the constant number of infectives for the remaining portion of time horizon. There is a fairly consistent decrease (approximately linear in appearance) in the number of infectives between $t = 1$ to $t = 38$, after an immediate drop in the first time step. After time $t = 38$, the number of infectives remains constant at 50.



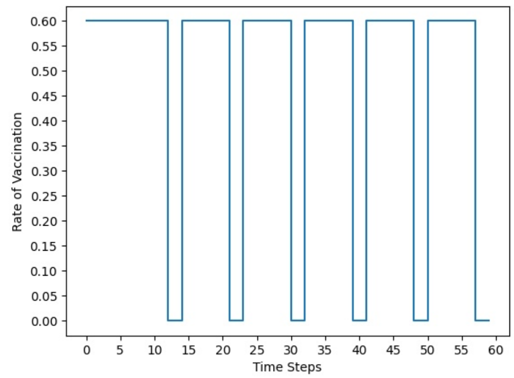
(a) Infectives **(b)** Rate of Vaccination

Figure 10. Binary Control Experiment with $\tau = 0.08$.

In **Figure 11**, τ is increased from 0.13 to 0.18. In this case, the rate at recovery of infectious individuals is higher, so more individuals are moving between I and S in each time step. Therefore, the need to intervene with vaccination is decreased. The algorithm administers the maximum vaccination rate for the first 12 time steps to substantially decrease the number of initial infectives. After the initial vaccination efforts, it adopts an oscillatory pattern to maintain control with each peak in infectives.



(a) Infectives



(b) Rate of Vaccination

Figure 11. Binary Control Experiment with $\tau = 0.18$.

CONCLUSIONS

A discrete-oriented approach to epidemiological modeling and optimization is presented. A continuous model for an Susceptible/Exposed/Infective/Removed/Susceptible (SEIRS) model was converted to a discrete representation in order to more accurately reflect real-life discrete values. Additionally, the discrete model accounted for a changing population at each time step. The algorithm utilized sought to find the minimum cost necessary to control an epidemic under a finite time horizon and discrete state space. Using python coding software, the algorithm was effectively carried out under various parameters and the numerical results were presented.

There are a couple of notable results from the completion of the research and numerical simulations. First, increasing the granularity of the vaccination rates helped to smooth out the graph of infectives versus time, so there are smaller significant peaks in the number of infectives, and the size of any rippling in the number of infectives is also minimized. It is possible that a larger peak becomes smaller and the rippling that occurs afterwards may increase slightly, as the number of infectives may be spread out throughout the duration of the time horizon. This was observed in the graph of the infectives for the 10-ary versus 20-ary control experiments. A second notable result occurred with changing the values of τ or γ. If the rate of death for the disease (γ) is decreased or the recovery rate (τ) is decreased, it became more difficult to control the epidemic. Therefore, the optimal solution was to administer the maximum rate of vaccination for the entire time horizon. If γ or τ is increased, it becomes easier to control the epidemic. In the case of an increased τ value, the number of individuals recovering increased, so there were less infectives at each time step to spread the disease; therefore, less vaccination efforts were required to control the epidemic. In the case of an increased γ value, the infection's increase in deadliness results in less need to vaccinate, since the disease kills more infectives at each time step; therefore, preventing the ability to transmit the disease to additional individuals.

ACKNOWLEDGEMENTS

The authors thank Madonna University's Department of Mathematics and Computer Science faculty and staff for their help and support. The authors also thank the reviewers of this work for their time and valuable feedback.

REFERENCES

1. Althaus, C.L. and Heng, K. (2020) The approximately universal shapes of epidemic curves in the Susceptible-Exposed-Infectious-Recovered (SEIR) model, *Scientic Reports*. *https://doi.org/10.1038/s41598-020-76563-8*
2. Arotaritei, D., Corciova, C., Constantin, G. (2018) Mathematical Models of Measles by Differential Equations in Virtual Education, *The International Scientific Conference eLearning and Software for Education* 3, 369–374.*https://doi.org/10.12753/2066-026X-18-197*
3. Blount, S. and Galambosi, A. and Yakowitz, S. (1997) Nonlinear and Dynamic Programming for Epidemic Intervention, *Applied Mathematics and Computation* 86, 123–136. *https://doi.org/10.1016/S0096-3003(96)00177-4*
4. Brauer, F. (2009) Mathematical Epidemiology is not an Oxymoron, *BMC Public Health* 9. *https://doi.org/10.1186/1471-2458-9-S1-S2*
5. Cruz-pacheco, G., Esteva, L., Vargas, C. (2014) Vaccination Strategies for SIR Vector-Transmitted Diseases, *BMB* 76, 2073–2090. *https://doi.org/10.1007/s11538-014-9999-6*
6. Daley, D.J. and Gani, J. and Yakowitz, S. (2000) An Epidemic with Individual Infectivities and Susceptibilities, *Mathematical and Computer Modeling* 32, 155–157.
7. Fan X, Wang L, Teng Z. (2016) Global dynamics for a class of discrete SEIRS epidemic models with general nonlinear incidence, *Advances in Difference Equations*. *https://doi.org/10.1186/s13662-016-0846-y*
8. Gerberry, D.J., Milner, F.A. (2009) An SEIQR model for childhood diseases, *JOMB* 59, 535–561. *https://doi.org/10.1007/s00285-008-0239-2*
9. Greer, M.L, Palin, K.A. (2012) Students in Differential Equations and Epidemiology Model a Campus Outbreak of pH1N1, *Journal of Microbiology and Biology Education* 13, 183–185. *http://dx.doi.org/10.1128/jmbe.v13.i2.429*
10. Hethcote, H.W. (2000) The Mathematics of Infectious Diseases, *SIAM Review* 42, 599–653. *https://doi.org/10.1137/S0036144500371907*
11. Jarynowski, A. (2013) Human-Human Interaction: Epidemiology, *arXiv*. *https://arxiv.org/ftp/arxiv/papers/1303/1303.7208.pdf*
12. Jia, J., Xiao, J. (2018) Stability analysis of a disease resistance SEIRS model with nonlinear incidence rate. *Adv Differ Equ* 75. *https://doi.org/10.1186/s13662-018-1494-1*
13. Ma, J. (2020) Estimating epidemic exponential growth rate and basic reproduction number, *Infectious Disease Modelling. https://doi.org/10.1016/j.idm.2019.12.009.*
14. Mummert, A. (2013) Studying the recovery procedure for the time-dependent transmission rate(s) in epidemic models, *JOMB* 67, 483–507. *https://www.doi.org/10.1007/s00285-012-0558-1*
15. Nierman, R. (2017) Combinatorial Approaches to Continuous Problems, Oakland University. *https://www.proquest.com/docview/1917681683*
16. Nierman, R. and Kruk, S. (2017) An Optimal Intervention Experiment with Multiple Controls, *AMO* 19. *https://camo.ici.ro/journal/vol19/v19a1.pdf*
17. Salpeter, E. and Salpeter, S. (1998) Mathematical Model for the Epidemiology of Tuberculosis, with Estimates of the Reproductive Number and Infection-Delay Function, *AJE* 147, 398–406. *https://doi.org/10.1093/oxfordjournals.aje.a009463*
18. Sethi, Suresh P., (1974) Quantitative Guidelines For Communicable Disease Control Program: A Complete Synthesis, *Biometrics* 30, *681–691. https://doi.org/10.2307/2529232*

ABOUT THE STUDENT AUTHOR

Tanner Snyder graduated from Madonna University in 2019 with a degree in Mathematics and a minor in Computer Science. He is currently working on his Masters Degree in Industrial Applied Mathematics at Oakland University with a focus in Discrete Mathematics.

PRESS SUMMARY

This research presents an optimal control model for epidemiology assuming a finite time period and discrete time steps. This research develops a Susceptible/Exposed/Infective/Removed/Susceptible (SEIRS) model and incorporates a changing population variable to account for births/deaths and immigration/emigration. A series of numerical results are presented using python software and graph generation. The research presented is of particular interest in the wake of the COVID-19 pandemic and can be applied to various SEIRS epidemics for research and modeling purposes.

Factors Associated with Surgery among South Asian American and Non-Hispanic White Women with Breast Cancer

Lydia Lo[*a] *& Jaya M. Satagopan*[b]

[a]*Department of Industrial and Systems Engineering, School of Engineering, Rutgers University, New Brunswick, NJ*
[b]*Department of Biostatistics and Epidemiology & Center for South Asian Quantitative Health and Education, School of Public Health, Rutgers University, New Brunswick, NJ*

https://doi.org/10.33697/ajur.2021.048

*Student: lpl33@scarletmail.rutgers.edu**
Mentor: satagopj@sph.rutgers.edu

ABSTRACT
South Asian American (SA) women are diagnosed with more aggressive breast cancer than non-Hispanic White (NHW) women. Understanding the factors associated with the types of surgery received by these women sheds light on disease management in these culturally distinct populations. We used data on age at diagnosis, stage, grade, estrogen and progesterone receptors, and surgery from 4,590 SA and 429,030 NHW breast cancer cases in the Surveillance, Epidemiology and End Results (SEER) program. We used logistic regression with surgery as the binary outcome (subcutaneous, total, or radical mastectomy (STRM) versus partial mastectomy, no, unknown or other (PNUM)) and included additive effects of all the variables and interactions of age, stage, grade, and estrogen and progesterone receptors with race/ethnicity. Type I error of 5% was used to assess statistical significance of the effects. SA were significantly more likely than NHW cases to receive STRM relative to PNUM surgery among women diagnosed at or after age 50 years and having localized stage disease (Odds Ratio (OR) = 1.27, 95% Confidence Interval (CI) = 1.06 – 1.52). Further, SA were significantly less likely than NHW cases to receive STRM relative to PNUM surgery among those diagnosed before age 50 years and having regional or distant stage disease (OR = 0.75, 95% CI = 0.59 – 0.95 for age at diagnosis < 40 years; OR = 0.77, 95% CI = 0.62 – 0.95 for age at diagnosis 40-49 years). The type of surgery received by SA and NHW women differ according to age at diagnosis and disease stage.

KEYWORDS
Breast Cancer; Surgery; Cancer Health Equity; Disease Characteristics; South Asian American; Non-Hispanic White; Logistic Regression; Interaction

INTRODUCTION

South Asians (SA) are individuals with origins in Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, and Sri Lanka. SA are a fast-growing minority population in the U.S, increasing by about 40% from around 3.9 million in 2010 to around 5.4 million in 2017.[1] SA women in the U.S confront a wide range of chronic diseases, including breast cancer. They are more likely than non-Hispanic White (NHW) women to be diagnosed with aggressive breast cancers characterized by younger age, regional or distant stage tumors, higher grade tumors, and estrogen receptor (ER)-negative tumors at the time of diagnosis.[2] There are many different surgical treatments available for breast cancer. The most common types of surgery are lumpectomy and mastectomy.[3, 4] Lumpectomy, also known as partial mastectomy, is a surgery to remove part of the breast that is affected by cancer.[5] There are various types of mastectomies – subcutaneous, total, or radical mastectomy – that involve removal of the entire breast, with or without the nipple, the areola, major muscles of the chest wall, and axillary or central lymph nodes.[6] In the U.S, early-stage breast cancers may be treated with lumpectomy, while large or later stage breast cancers may be removed by mastectomy.[7] In contrast, the rate of mastectomy relative to lumpectomy for breast cancers is high in Asian countries, regardless of disease stage.[8] Little is known about surgical treatment for breast cancer among SA women in the U.S, whether SA women are more likely than NHW women to undergo mastectomy, and whether any differences in surgery between SA and NHW persists after adjusting for tumor characteristics. To address this gap, we undertook this study to examine factors associated with type of breast cancer surgery in SA and NHW women using data from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program.[9] Understanding whether mastectomy rates are high among SA women in the US even after accounting for tumor characteristics such as stage can provide valuable insights into breast cancer-related treatment decisions in this culturally distinct minority population.

METHODS AND PROCEDURES

Data Source

We extracted breast cancer data from 18 population-based SEER registries that cover about 28% of the U.S. population[9]. We completed a data use agreement with SEER and downloaded a case listing of female breast cancers diagnosed between the years 2000 and 2016 via the SEER*Stat software package.[9] Using the International Statistical Classification of Diseases (ICD), we abstracted data corresponding to ICD code "ICD-O-3: Breast from ages 0 to 85+". This ICD classification includes the following breast cancers: 8500/3: Infiltrating duct carcinoma, NOS; 8521/3: Infiltrating ductular carcinoma; 8522/3: Infiltrating duct and lobular carcinoma; 8523/3: Infiltrating duct mixed with other types of carcinoma; and 8524/3: Infiltrating lobular mixed with other types of carcinoma, 8540/3: Paget disease, mammary; 8541/3: Paget disease and infiltrating ductal carcinoma of breast; and 8543/3: Paget disease and intraductal carcinoma.

Outcome Variable

We obtained surgery information from the SEER variable "RX Summ—Surg Prim Site (1998)+" and used SEER surgery codes[10] to categorize this variable as subcutaneous mastectomy (surgery code 30), total mastectomy (surgery codes 40-49 and 75), radical mastectomy (surgery codes 50-74), partial mastectomy (surgery codes 20-24), no surgery (code 0), other surgery (codes 19 or 90), and unknown if surgery performed since patient information is from death certificate only (code 99). The binary outcome in our study was surgery, categorized as subcutaneous, total, or radical mastectomy (abbreviated as STRM) and partial mastectomy, no, other or unknown surgery (abbreviated as PNUM).

Explanatory Variables

The explanatory variables were age at diagnosis, race/ethnicity, tumor stage, tumor grade, estrogen receptor (ER) status, and progesterone receptor (PR) status. We used the SEER variable "Age recode with < 1 year olds" to obtain age at diagnosis of breast cancer. We used the SEER variable "Race/ethnicity" to obtain data from non-Hispanic White (NHW), Asian Indian or Pakistani, NOS (1988+), Asian Indian (2010+), and Pakistani (2010+) women. The latter three race/ethnicity consisting of women of Asian Indian or Pakistani origin were grouped as South Asian (SA). We obtained tumor stage using SEER variables "SEER Summary Stage 1977", "SEER Summary Stage 2000" and "SEER Combined Summary Stage 2000" for cases diagnosed in 2000, from 2001 to 2003, and from 2004 to 2016, respectively, and categorized stage as localized, regional, distant, and unknown. SEER variable "Grade" provides tumor grade as I, II, III, IV, or unknown. SEER variables "ER Status Recode Breast Cancer (1990+)" and "PR Status Recode Breast Cancer (1990+)" provide ER and PR statuses, respectively, as positive, negative, borderline, or unknown. Patients with unknown stage or grade and unknown or borderline ER or PR statuses were removed from analyses.

Statistical Analysis

We calculated mean and standard deviation (SD) to examine baseline characteristics of age at diagnosis and calculated frequencies and proportions to examine baseline characteristics of race/ethnicity, stage, grade, ER status, PR status and surgery. We obtained boxplot of age at diagnosis of breast cancer of SA and NHW women receiving STRM and PNUM surgeries. We used a two-sample t-test to compare the mean age at diagnosis of breast cancer between SA and NHW women, and chi-squared tests to compare frequencies of race/ethnicity, stage, grade, ER status and PR status between SA and NHW women. Prior to conducting further analyses, we categorized age at diagnosis as "< 40 years", "40 – 49 years", and "≥ 50 years". These age categories have been previously used in breast cancer studies and are taken to denote early-onset disease (age < 40 years), onset during perimenopausal period (age 40-49 years), and onset during post-menopausal period (age ≥ 50 years).[2] We also created binary variables for stage (localized versus regional or distant) and grade (I and II versus III and IV). Table 1 shows all the variables and their categories used in our analyses.

Using these categorized variables, we conducted logistic regression analyses to evaluate the association between age at diagnosis, stage, grade, ER and PR statuses and surgery.[11] First, we conducted univariable logistic regression analyses to gain insights into relationships between each explanatory variable and surgery type without adjusting for other explanatory variables. Next, we fitted two multivariable logistic regression models – first by including only additive effects of all the explanatory variables, and next by including all the additive effects and pairwise multiplicative interactions between race/ethnicity and age, stage, grade, ER status, and PR status. We calculated odds ratios (OR) and 95% confidence intervals (CI) and tested the statistical significance of the odds ratios using Wald tests. In the multivariable model with interaction terms, we derived odds ratios for race/ethnicity in distinct categories of age at diagnosis and stage as shown in the Appendix. All hypothesis tests were conducted at 5% type I error.

	Binary Categories	
Variables in analysis	**Category 0 (or baseline category)**	**Category 1 (or comparison category)**
Outcome Variable		
Surgery (mastectomy type)	Partial, None, Unknown, Other (PNUM)	Subcutaneous, Total, Radical (STRM)
Explanatory Variables		
Age		
Age 40 years	Age ≥ 50 years	Age < 40 years
Age 40-49 years	Age ≥ 50 years	Age 40-49 years
Race	Non-Hispanic White	South Asian
Grade	I and II	III and IV
Stage	Localized	Regional and Distant
ER	Positive	Negative
PR	Positive	Negative

Table 1. Variables used in the analysis and their binary categories. Note that since age has 3 categories, we created two dummy variables corresponding to the categories Age < 40 years and Age 40-49 years, keeping Age ≥ 50 years as the baseline category.

When the 95% CI for an OR corresponding to a variable included the value 1, the corresponding p-value for the Wald test also exceeded 0.05 as expected, suggesting no statistically significant association between the variable and surgery type or suggesting that individuals from distinct categories of that variable were equally likely to receive STRM relative to PNUM surgery at the 5% significance level. A statistically significant odds ratio for an interaction between race/ethnicity and another variable was taken as an indication that the effect of race/ethnicity on type of surgery differed across the categories of that variable.

The multivariable models fitted were evaluated using goodness-of-fit measures such as Akaike information criterion (AIC), Bayes information criterion (BIC), and area under the receiver operating characteristic curve (AUC). The AIC is defined as negative two times the maximum log-likelihood plus two times the number of parameters in the model. The BIC is defined as two times the maximum log-likelihood subtracted by the number of parameters in the model times the log of the sample size. The AUC is a measure of how well the model is able to distinguish between the binary surgery types. The best fitting logistic regression model among those considered will ideally have the smallest AIC and BIC and the largest AUC.

All statistical analyses were conducted using the R programming language (version 3.6.3).[12]

RESULTS

We obtained data for a total of 449,040 SA and NHW female breast cancer cases diagnosed between 2000 and 2016 from SEER. After removing cases with unknown stage, unknown grade, and unknown or borderline ER and PR statuses, the dataset for our analysis included 4,590 SA and 429,030 NHW cases. Table 2 shows the patients' characteristics. Figure 1 shows boxplots of age at diagnosis of breast cancer in SA and NHW cases receiving STRM and PNUM surgery. SA cases were more likely to be younger than NHW cases among women receiving STRM surgery (mean (SD) age at diagnosis = 52.6 (13.6) years for SA and 59 (14.3) years for NHW; $p < 0.05$) and among those receiving PNUM surgery (mean (SD) age at diagnosis = 55.2 (12.7) years for SA and 62.4 (13.0) years for NHW; $p < 0.05$). Further, in both STRM and PNUM surgery groups, SA cases were significantly more likely to have regional or distant stage, higher grade, ER negative and PR negative tumors than NHW cases ($p < 0.05$ for all comparisons).

Table 3 provides the results of univariable and multivariable logistic regression analyses. In univariable analyses (column 3 of Table 3), the characteristics of cases that were significantly more likely to receive STRM surgery relative to PNUM surgery included younger age, regional or distant stage, higher grade, ER negative, and PR negative tumors. Further, SA cases were significantly more likely than NHW cases to receive STRM relative to PNUM surgery (OR = 1.25, 95% CI = 1.18 – 1.33).

In all the multivariable analyses considered, cases with younger age at diagnosis, regional or distant stage, higher grade, ER negative status and PR negative status were significantly more likely to receive STRM relative to PNUM surgery. In multivariable analyses that included only additive effects (column 4 of Table 3), there was no statistically significant association between race/ethnicity and surgery type: both SA and NHW cases were equally likely to receive STRM relative to PNUM surgery (OR = 1.03, 95% CI = 0.97, 1.10). In contrast, in multivariable analyses that included additive and interaction effects (column 5 of Table 3), race/ethnicity had a statistically significant interaction with age at diagnosis and stage, thus highlighting the critical role of these variables on the relationship between race/ethnicity and surgery type. Table 4 gives the odds ratios for race/ethnicity in distinct categories of age at diagnosis and stage. Among cases with localized stage disease diagnosed before age 40 years and between age

40-49 years, SA and NHW women were equally likely to receive STRM relative to PNUM surgery (OR = 0.94, 95% CI = 0.77 – 1.14 for age at diagnosis < 40 years; OR = 0.96, 95% CI = 0.80 – 1.15 for age at diagnosis 40-49 years). However, among cases with localized disease diagnosed at age 50 years or later, SA women were significantly more likely than NHW women to receive STRM relative to PNUM surgery (OR = 1.33, 95% CI = 1.19 – 1.48). In contrast, among cases with regional or distant stage disease diagnosed before age 40 years and between age 40-49 years, SA women were significantly less likely than NHW women to receive STRM relative to PNUM surgery (OR = 0.79, 95% CI = 0.63 – 0.98 for age at diagnosis < 40 years; OR = 0.80, 95% CI = 0.69 – 0.93 for age at diagnosis 40-49 years). However, among cases with regional or distant stage disease diagnosed at age 50 years or later, SA women and NHW women were equally likely to receive STRM relative to PNUM surgery (OR = 1.12, 95% CI = 0.98 – 1.27).

	STRM Surgery				**PNUM Surgery**			
Variable	**NHW (N = 159,533)**		**SA (N = 1,956)**		**NHW (N = 269,497)**		**SA (N = 2,634)**	
Age	**Mean**: 59.03	**SD**: 14.26	**Mean**: 52.63	**SD**:13.53	**Mean**: 62.36	**SD**: 12.99	**Mean**: 55.17	**SD**: 12.74
Grade								
Grade I	23,944		182		66,523		455	
Grade II	65,713		744		114,995		1,036	
Grade III	62,959		929		76,800		1004	
Grade IV	1,513		17		1,684		16	
Stage								
Localized	82,152		930		195,244		1,708	
Regional	72,428		961		59,763		722	
Distant	4,626		60		12,608		174	
ER Status								
Positive	124,042		1,445		226,104		2,100	
Negative	35,491		511		43,393		534	
PR Status								
Positive	106,566		1,255		198,461		1,870	
Negative	52,967		701		71,036		764	

Table 2. Median and standard deviation (SD) of age at diagnosis of breast cancer and frequencies of grade, stage, ER status and PR status for SA and NHW breast cancer patients receiving STRM and PNUM surgery.

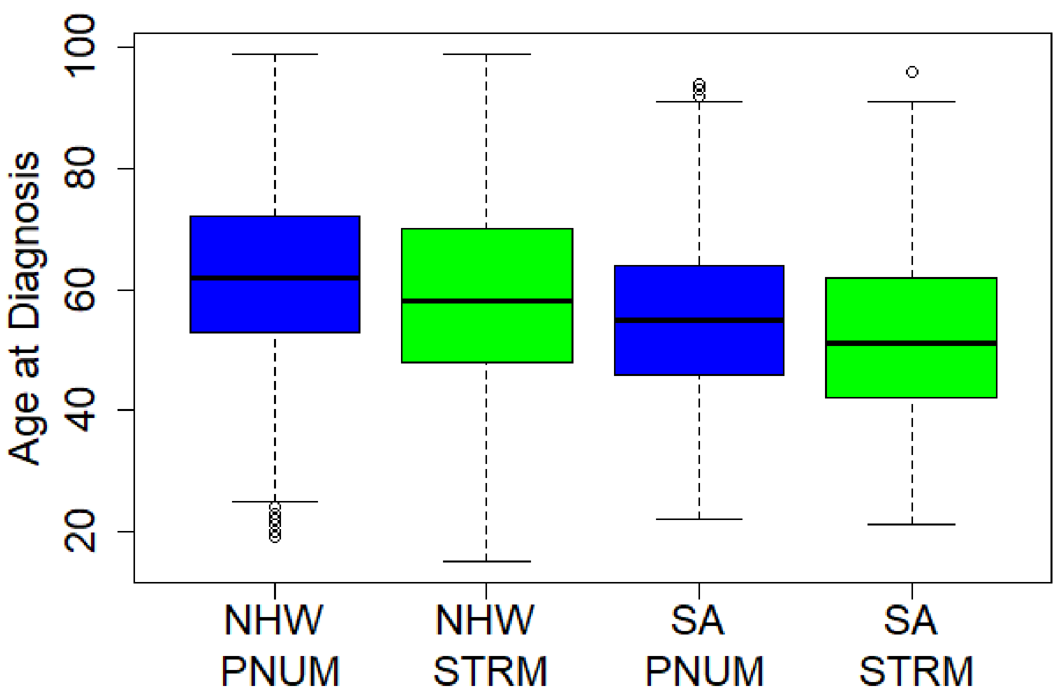


Figure 1. Boxplot of age at diagnosis of breast cancer of non-Hispanic White (NHW) and South Asian (SA) cases receiving PNUM and STRM surgery.

The multivariable model with additive and interaction terms had the smallest AIC but not the smallest BIC (Table 3). It had the same AUC as the multivariable additive model. As a benchmark for comparison, a multivariable additive model without race/ethnicity had AIC of 541585, BIC of 541661.8, and AUC of 0.6516 which are considerably worse than models containing race/ethnicity (detailed result of latter model is not shown).

Variable	Category	Odds Ratio (95% CI)		
			Multivariable Analysis	
		Univariable Analysis	Model with additive effects only	Model with Additive and interaction effects
Additive effects				
Race/ethnicity	SA	1.25 (1.18, 1.33)	1.03 (0.97, 1.10)	1.33 (1.19, 1.48)
	NHW	1.0	1.0	1.0
Age	<40 years	2.56 (2.49, 2.64)	2.34 (2.27, 2.41)	2.36 (2.28, 2.43)
	40-49 years	1.55 (1.52, 1.57)	1.53 (1.51, 1.56)	1.54 (1.51, 1.57)
	≥ 50 years	1.0	1.0	1.0
Stage	Regional or distant	2.51 (2.48, 2.54)	2.39 (2.35, 2.42)	2.39 (2.36, 2.42)
	Localized	1.0	1.0	1.0
Grade	III or IV	1.66 (1.64, 1.68)	1.28 (1.26, 1.30)	1.28 (1.26, 1.30)
	I or II	1.0	1.0	1.0
ER status	Negative	1.49 (1.47, 1.51)	1.07 (1.05, 1.10)	1.07 (1.05, 1.10)
	Positive	1.0	1.0	1.0
PR status	Negative	1.39 (1.37, 1.41)	1.15 (1.12, 1.17)	1.15 (1.12, 1.17)
	Positive	1.0	1.0	1.0
Interaction effects				
Race/ethnicity * Age	SA * Age < 40 years	-	-	0.70 (0.58, 0.85)
	SA * Age 40-49 years	-	-	0.72 (0.62, 0.83)
Rage * Stage	SA * regional or distant	-	-	0.84 (0.74, 0.95)
Race * Grade	SA * Grade III or IV	-	-	0.94 (0.82, 1.08)
Race * ER	SA * ER negative	-	-	0.96 (0.77, 1.19)
Race * PR	SA * PR negative	-	-	0.99 (0.82, 1.21)
AIC			521985.5	521955.6
BIC			522073	522108.7
AUC			0.6524	0.6524

Table 3. Results of univariable and multivariable logistic regression analyses.

Age	Stage	Odds Ratio (OR)	95% Confidence Interval (CI)
< 40 years	Localized	0.94	0.77, 1.14
	Regional / Distant	0.79	0.63, 0.98
40-49 years	Localized	0.96	0.80, 1.15
	Regional / Distant	0.80	0.69, 0.93
≥ 50 years	Localized	1.33	1.19, 1.48
	Regional / Distant	1.12	0.98, 1.27

Table 4. Odds ratio for receiving STRM relative to PNUM surgery for SA race/ethnicity compared to NHW race/ethnicity in distinct categories of age and stage, and 95% confidence interval (CI). These results are based on multivariable logistic regression model with additive and interaction terms,

DISCUSSION

Surgery is one of the treatments available for breast cancer. Lumpectomy or partial mastectomy and mastectomy are the common surgical treatments for breast cancer. Mastectomy is more common than lumpectomy in Asian countries, regardless of disease stage. However, little is known about the use of mastectomy versus lumpectomy among breast cancer cases of SA race/ethnicity, who are among the fast-growing minority populations in the U.S, and how this compares to the use of surgeries among NHW breast cancer cases. Therefore, in this study, we examined factors associated with type of surgery (STRM versus PNUM) among female SA versus NHW breast cancer cases using data from SEER. Our analyses showed that: (i) there are statistically significant differences between SA and NHW breast cancer cases in the use of STRM relative to PNUM surgery; and (ii) these differences vary according to age at diagnosis and stage.

Our analyses showed the importance of interactions of race/ethnicity with age and stage. While the model with interactions had the smallest AIC, its BIC was larger than that of an additive model. This could be because BIC penalizes for having more terms in the model.

There can be several reasons why SA are more likely than NHW to have STRM relative to PNUM surgery when their disease is diagnosed at or after age 50 and they have localized stage disease but are less likely than NHW to have STRM relative to PNUM surgery when their disease is diagnosed before age 50 and they have regional or distant stage disease. Cultural factors are known to play a role in treatment preferences of female Chinese American breast cancer patients.[13] Health decisions of South Asian women living in Canada are strongly influenced by family and community responsibilities.[14] Further research is needed to examine cultural aspects of treatment preferences among South Asian American women, *i.e.*, South Asian women living in the United States. Whether other factors such as comorbidities play a role in surgery decisions of SA women is also an important question that requires further research.

Our work has several strengths and limitations. A strength of our study is the significance of our research question since SA are a rapidly growing minority population in the U.S and, hence, understanding their breast cancer health and treatment decisions is important. Another strength of our study is the availability of data on a large number of SA and NHW cases from SEER.

Several limitations must be considered when evaluating the results of our study. The registries included in the SEER 18 database cover only 28% of the entire US population. There is a large SA population in other major cities such as Chicago, Dallas, Houston, New York, and Washington DC that are not included in the databases that we have used in this study. The SA cases in our study are of Asian Indian and Pakistani origin. Cases with ancestry from other SA countries – Bangladesh, Bhutan, Maldives, Nepal, and Sri Lanka – are not included in our study since their race/ethnicity are not available through SEER. Our analyses are not adjusted for comorbidities since these data are not available through SEER. A well-designed study of SA and NHW cases will be required to address these limitations.

CONCLUSIONS

Our analyses illustrated the important role of interaction terms in logistic regression model. An additive multivariable logistic regression model showed that SA and NHW cases were equally likely to receive STRM relative to PNUM surgery. Including interactions between race/ethnicity and age as well as stage revealed that differences in the type of surgery between SA and NHW varied according to age and stage categories. Disease characteristics play a key role on the type of surgery received by SA and NHW breast cancer patients. Further research is needed to gain insights into reasons underlying the choice of surgery made by SA women with specific disease characteristics.

REFERENCES

1. South Asian Americans Leading Together (n.d) Resources and Fact Sheets. Retrieved January 6, 2021, from *https://saalt.org/resources/resources-factsheets/*

2. Satagopan, J. M., Stroup, A., Kinney, A. Y., Dharamdasani, T., Ganesan, S., & Bandera, E. V. (2021) Breast cancer among Asian Indian and Pakistani Americans: A surveillance, epidemiology and end results-based study. *International journal of cancer*, *148*(7), 1598–1607. *https://doi.org/10.1002/ijc.33331*

3. Invasive Ductal Carcinoma: Diagnosis, Treatment, and More. (2019) Retrieved January 6, 2021, from *https://www.breastcancer.org/symptoms/types/idc*

4. Mayo Clinic. (2019) Paget's disease of the breast. Retrieved January 6, 2021, from *https://www.mayoclinic.org/diseases-conditions/pagets-disease-of-the-breast/diagnosis-treatment/drc-20351084*

5. Mayo Clinic. (2018) Lumpectomy. Retrieved January 6, 2021, from *https://www.mayoclinic.org/tests-procedures/lumpectomy/about/pac-20394650*
6. Oyelowo, T. (2009) Breast Conditions. Retrieved January 6, 2021, from *https://www.sciencedirect.com/science/article/pii/B9780323046015500083*
7. Mayo Foundation for Medical Education and Research. (2020) *Breast cancer surgery*. Mayo Clinic. *https://www.mayoclinic.org/tests-procedures/breast-cancer-surgery/about/pac-20385255*
8. Sinnadurai, S., Kwong, A., Hartman, M., Tan, E. Y., Bhoo-Pathy, N. T., Dahlui, M., See, M. H., Yip, C. H., Taib, N. A., & Bhoo-Pathy, N. (2018) Breast-conserving surgery *versus* mastectomy in young women with breast cancer in Asian settings. *BJS open*, *3*(1), 48–55. *https://doi.org/10.1002/bjs5.50111*
9. Surveillance, Epidemiology, and End Results (SEER) Program (*www.seer.cancer.gov*) SEER*Stat Database: Incidence - SEER 18 Regs Custom Data (with additional treatment fields), Nov 2018 Sub (2000-2016) <Katrina/Rita Population Adjustment> - Linked To County Attributes - Total U.S., 1969-2017 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, released April 2019, based on the November 2018 submission.
10. Adamo MB, Johnson CH, Ruhl JL, Dickie, LA, (Eds.). 2013 SEER Program Coding and Staging Manual. National Cancer Institute, NIH Publication number 13-5581, Bethesda, MD
11. Agresti A. Categorical data analysis. John Wiley & Sons; 2003 Mar 31.
12. R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL *https://www.R-project.org/*
13. Killoran, M., & Moyer, A. (2006) Surgical treatment preferences in Chinese-American women with early-stage breast cancer. *Psycho-Oncology: Journal of the Psychological, Social and Behavioral Dimensions of Cancer, 15(11),* 969–984. *https://doi.org/10.1002/pon.1032*
14. Bottorff, J. L., Johnson, J. L., Bhagat, R., Grewal, S., Balneaves, L. G., Clarke, H., & Hilton, B. A. (1998) Beliefs related to breast health practices: The perceptions of South Asian women living in Canada. *Social Science & Medicine, 47(12),* 2075-2085. *https://doi.org/10.1016/S0277-9536(98)00346-3*

ABOUT STUDENT AUTHOR

Lydia Lo graduated cum laude with a Bachelor of Science degree in Industrial Engineering from Rutgers University in May 2021. She currently works as an Operations Industrial Engineer at the United States Postal Service.

PRESS SUMMARY

South Asians (SA) are a fast-growing minority population in the U.S and confront a wide range of chronic disease, including breast cancer. SA women are diagnosed with more aggressive breast cancer than non-Hispanic White (NHW) women, characterized by younger age at diagnosis, regional or distant stage, higher grade, and ER negative tumors. Surgery – especially lumpectomy or mastectomy – is one of the treatments available for breast cancer. In the U.S, lumpectomy is used to treat early-stage breast cancers, while mastectomy is used for later stage breast cancers. In contrast, mastectomy is more commonly used in Asian countries, regardless of disease stage. It is important to understand the type of surgery and the factors that contribute to determining the type of surgery in SA breast cancer cases in the U.S relative to NHW cases. Our analyses reveal differences in the type of surgery received by SA and NHW breast cancer cases, and these differences vary according to age at diagnosis and disease stage. Further research is needed to gain deeper insights into these differences.

APPENDIX

The Results section of the paper includes odds ratios corresponding to race/ethnicity in distinct subgroups of age and stage. The approach we used to obtain these odds ratios are given below.

Denote π as the probability that a person receives STRM surgery, given age, race/ethnicity, stage, grade, ER and PR. The logistic regression model consisting of interaction of race with age and stage is:

$$\log\{ \pi / (1 - \pi) \} = b_0 + b_1{*}x_1 + b_2{*}x_2 + b_3{*}race + b_4{*}stage + b_5{*}grade + b_6{*}er + b_7{*}pr + d_1{*}x_1{*}race + d_2{*}x_2{*}race + d_3{*}stage{*}race$$

The left hand of this model is known as the natural logarithm of odds or log-odds. The explanatory variables on the right hand side are as follows. First, x_1 and x_2 are dummy variables corresponding to age at diagnosis such that $x_1 = 1$ if a person's age at diagnosis is < 40 years and $x_1 = 0$ otherwise, and $x_2 = 1$ if a person's age at diagnosis is between 40 and 49 years and $x_2 = 0$ otherwise. Thus, $x_1 = 0$ and $x_2 = 0$ for person whose age at diagnosis is ≥ 50 years. Further, stage, grade, er, pr, and race are binary variables coded as given in Table 1 of the manuscript. The model parameters are as follows: b_0 is the y-intercept, b_i (b_1, b_2, b_3, etc.) are the additive effects of the explanatory variables and d_i (d_1, d_2, d_3) are the effects of the interaction of race with age at diagnosis categories and stage, respectively. Although our multivariable analysis reported in the Results section includes interaction of race with grade, ER and PR, we have not included these interactions in writing the above model for simplicity of exposition. [Further, the estimated parameters for interactions of race grade, ER and PR are very small (near 0) and are not statistically significant in our analyses. Hence, including these interaction terms did not impact our estimated odds ratios corresponding to race in distinct categories of age and stage]

Odds ratio corresponding to SA relative to NHW race among those diagnosed at age ≥ 50 years and having localized stage disease and 95% CI:

Consider a person of age at diagnosis ≥ 50 years with localized stage disease and NHW race. For this person, $x_1 = 0$, $x_2 = 0$, race = 0 and stage = 0. Plugging these values into the above model, the log-odds for this person will be: $b_0 + b_5{*}grade + b_6{*}er + b_7{*}pr$. Consider a person of age at diagnosis ≥ 50 years with localized stage disease and SA race. For this person, $x_1 = 0$, $x_2 = 0$, race = 1 and stage = 0. The log-odds for this person will be: $b_0 + b_3 + b_5{*}grade + b_6{*}er + b_7{*}pr$.

To get the odds ratio of STRM versus PNUM surgery in SA race relative to NHW race among those of age at diagnosis ≥ 50 years and localized stage disease, take the difference between the two log odds and then the exponent of that. The difference between the two log odds is simply b_3. Hence, the odds ratio is $\exp\{ b_3 \}$.

The standard error corresponding to the estimated value of b_3 can be obtained from the covariance matrix output of the logistic regression model in the R programming language. Suppose we denote this as SE_3. The 95% CI for the above odds ratio can be approximated as $\exp\{ b_3 +/- 1.96 * SE_3 \}$.

Odds ratio corresponding to SA relative to NHW race among those diagnosed at age ≥ 50 years and having regional or distant stage disease and 95% CI:

Consider a person of age at diagnosis ≥ 50 years with regional or distant stage disease and SA race. For this person, $x_1 = 0$, $x_2 = 0$, race = 0 and stage = 1. The log-odds for this person will be: $b_0 + b_4 + b_5{*}grade + b_6{*}er + b_7{*}pr$. Consider a person of age at diagnosis ≥ 50 years with regional or distant stage disease and SA race. For this person, $x_1 = 0$, $x_2 = 0$, race = 1 and stage = 1. The log-odds for this person will be: $b_0 + b_3 + b_4 + b_5{*}grade + b_6{*}er + b_7{*}pr + d_3$.

The difference between the two log odds is $b_3 + d_3$. Thus, the odds ratio for STRM relative to PNUM surgery in SA compared to NHW women with age at diagnosis ≥ 50 years and having regional or distant stage disease is $\exp\{ b_3 + d_3 \}$.

To calculate a 95% CI for this odds ratio, we need the standard error of b_3 + d_3. We shall denote this standard error as SE_{34}. It is the square root of the variance i.e., SE_{34} is the square root of the variance $Var(b_3 + b_4)$, which is given by $Var(b_3) + Var(d_4) + 2*Cov(b_3, d_4)$. These variances and covariances can be obtained from the covariance matrix output of the logistic regression model in the R programming language. The 95% CI can then be approximated as: $\exp\{ b_3 + d_4 +/- 1.96 * SE_{34} \}$.

The odds ratios can be written in this manner for other categories of age and stage.

The Effect of Perceived Uncertainty on Competitive Behavior

Duncan James Drewry[*a] *& Zachary Reese*[b]

[a]*Weinberg Institute for Cognitive Science, University of Michigan, Ann Arbor, MI*
[b]*Department of Psychology, University of Michigan, Ann Arbor, MI*

https://doi.org/10.33697/ajur.2021.049

*Student: ddrewry@umich.edu**
*Mentors: zreese@umich.edu**

ABSTRACT
How do people behave in the face of uncertainty? Some studies suggest that even when they are unaware of how others will behave, people default to cooperative behavior; however, other research suggests that uncertainty leads to more competitive behavior. Little research has examined how individual differences moderate such behavioral decisions. This study proposes that a stable (dispositional) sense of justice may, ironically, lead to more competitive behavior. Specifically, people who score highly in belief in a just world, system justification, and religiosity, and low in ambiguity tolerance may be more inclined to compete rather than cooperate because they believe people who experience positive outcomes deserve those outcomes regardless of the means taken to achieve them. Across two studies, participants (*N* = 288) engaged in a prisoner's dilemma game — a task where they must choose to compete or cooperate — and completed the aforementioned individual difference measures. Results show that people tended to cooperate, but those high in system justification and belief in a just world were more likely to compete. In other words, people with a strong sense of cosmic justice were likely to exhibit competitive behavior under uncertain conditions.

KEYWORDS
Ambiguity Tolerance; Competition; Cooperation; Just World Beliefs; Prisoner's Dilemma; Prosocial Behavior; Religiosity; System Justification; Uncertainty

INTRODUCTION

With the rise of the COVID-19 pandemic, economic turmoil, and the threat of international conflict, our society lives with a great amount of daily uncertainty. Amidst this uncertainty, people must decide how to interact with one another. Many choose to compete (*e.g.,* hoarding groceries and toilet paper). Others choose to cooperate (*e.g.,* socially distancing and wearing masks). Why do some choose one strategy over the other? Might one's sense of cosmic justice impact their social behavior in times of uncertainty?

Uncertainty, in the sense that individuals have no information about how others will behave, is a complex issue that has been tackled across fields, including philosophy, computer science, economics, business management, sociology, and psychology.[1–9] Research into uncertainty's effect on human behavior suggests a lack of consensus. Prior studies — although limited in number[10] — have shown a dialectical struggle between cooperation and competition. Psychological research has yet to explore specifically how such 'cosmic uncertainty' — as defined in this study as the general feeling that worldly events are random, lacking order or fairness — affects social decision-making. Individuals who believe in randomness believe that good life behaviors may not result in good life outcomes. For instance, a person who donates money to charity is no more or less likely to contract a terminal illness than someone who does not donate to charity. Conversely, 'cosmic justice' is operationally defined as the general feeling that the world is a just, fair place where people generally get what they deserve. People who believe in cosmic justice are more inclined to believe that good things happen to good people and bad things happen to bad people. Thus, a person who donates money to charity may have a lower probability of contracting a terminal illness than someone who does not donate to charity. There is no single scale assessing a sense of cosmic justice, but there are many adjacent scales which will be introduced in the "Individual

Differences" section below. First, it is worth examining people's general tendencies toward cooperative versus competitive behavior when situations do not explicitly call for one behavior or the other.

Uncertainty and Cooperation

Some studies suggest that people performing tasks with others default to cooperative behavior.[1, 2] This means that humanity's altruistic tendencies tend to outweigh situational uncertainty. One reason to default to altruism is the concept of evolutionary fitness.[11] Evolutionary psychology delves into how human and animal behavior tends to maximize the agent's evolutionary fitness — the ability to pass on our gene pool in the form of offspring. Our human ancestors were more likely to survive and pass along their genes if they cooperated with those who they are related to — such as offspring and family members. Early humans faced uncertainty; by working together, as opposed to competing with each other, they were able to hunt large game that could feed all.

There are, too, more practical purposes for cooperation. For example, people factor their reputation into their competitive or cooperative decision-making processes.[12] When an individual is pitted against another person in multiple social situations, they tend to avoid giving negative impressions of themselves so as not to harm any potential future relationship. Indeed, in studies of game theory, participants tend toward cooperation and mutual respect — especially when they are aware that they will be playing multiple games in succession with that same partner.[12] Thus, there is reason to believe that, under conditions of uncertainty, human behavior tends to favor cooperation over competition.

Uncertainty and Competition

However, other evidence suggests that uncertainty leads to more displays of competitive behavior. For instance, students in the classroom who are ranked against each other tend to withhold assistance from their peers.[3] This competitive effect was seen when personal and peer rankings were unknown suggesting that uncertainty plays a significant role in the decision-making process. When resources are scarce and people do not know their likelihood of attaining those resources, they hold less regard for others.[13]

Similarly, fellow employees are more likely to compete rather than cooperate when payoffs are uncertain.[14] Employees at all levels in a performance-based compensation scheme who did not know how many 'winners' would be named competed against each other at a much higher rate. This willingness to enter into competitive behavior was strictly determined by the ambiguous payoff conditions for the employees.

Thus, the effect uncertainty has on social decision-making is not fully understood[15] as researchers have arrived at competing conclusions.

Individual Differences in World Beliefs

To predict whether people will compete or cooperate in uncertain conditions, it may be helpful to examine the world beliefs people tend to carry into situations. In other words, do people generally believe in randomness or cosmic justice? The purpose of this study is to see whether individual differences in justice beliefs predict behavioral outcomes, namely cooperation or competition. Such individual differences include belief in a just world,[16] ambiguity tolerance,[17] and system justification.[18] Belief in a just world is one's tendency to believe that people tend to get what they deserve.[16] Ambiguity tolerance is defined as a measure of one's preference and comfortability with uncertain conditions.[17] Similarly to just world beliefs, system justification measures how well individuals believe the current system — either cultural, societal, organizational, or economic — is inherently justified.[18] Each of these individual differences relates to beliefs in cosmic justice. Specifically, people who score highly in belief in a just world and system justification and low in ambiguity tolerance would be more inclined to believe that good things tend to happen to good people.

Current Study

This study proposes that people who have a stronger inherent sense of justice may engage in more competitive (versus cooperative) behavior. Whereas cooperation tends to benefit the group, competition tends to benefit the individual. So, people with a strong sense of justice may feel entitled to the individual spoils of competition and attribute their positive outcomes to cosmic justice rather than their individual decision to compete. To examine this hypothesis, participants in this study completed a prisoner's dilemma game. The prisoner's dilemma game is a two-player game in which each player can choose to either cooperate or compete.[19] However, the game is played without knowing what one's partner selected, and outcomes are mutually determined such that competing tends to benefit the individual and cooperating tends to benefit the group. Thus, the prisoner's dilemma game is a perfect vehicle to study competition and cooperation when outcomes are uncertain.

Participants in this study attended virtual game sessions where they played a simple prisoner's dilemma game with others they met at the sessions. Sessions were designed to maximize participants' uncertainty as to how others would behave. Two studies were conducted. This exploratory first study aimed to understand whether and how people's belief systems predicted their competitive versus cooperative behavior. The second study aimed to experimentally manipulate conditions of cosmic uncertainty by priming participants with a fair or unfair condition. Additionally, a religiosity index was added to the second study as people who are highly religious generally have a strong sense of cosmic justice. Thus, it may show similar or related results to the original three scales in situations of competition.

METHODS AND PROCEDURES

Study 1

Participants

Students were recruited from introductory psychology courses (N = 151) at a large midwestern university for course credit. Though demographic information was not collected for this sample, here demographics data from the subject pool at large is reported and from our own past studies. Subject pool participants are generally 61% Female, 59.6% White, and average 18.9 years old. Participants signed up for times resulting in groups of no less than three and no more than ten. Use of this sample was authorized by the authors' Institutional Review Board who deemed the current line of research exempt (HUM00189456).

Materials

Four constructs were used to create a Qualtrics survey: Belief in a Just World,[16] Ambiguity Tolerance,[17] System Justification,[18] and the Positive and Negative Affect Scale (PANAS)[20] (the PANAS was included as a distractor). Higher scores on the belief in a just world scale predict important attitudes such as a stronger belief in an individual's own ability to affect change on the world based on their actions.[16] A high ambiguity tolerance score predicts that individuals will manage well in unfamiliar, ambiguous, or cross-cultural situations.[17] Higher scores on the system justification scale predict behaviors that uphold the current societal system.[18] Internal reliability as measured by Cronbach's alpha was high for system justification (α = .82) but low for belief in a just world (α = .69) and ambiguity tolerance (α = .61).

A task followed where written instructions asked participants to count specific letters (e.g., "e" and "g") on a page from a textbook. This task was also used in the prisoner's dilemma game as the assignment they would either split with (cooperate) or assign to (compete) their partner in the game.

Design

The independent variables were the three scales measuring beliefs about uncertainty, and the dependent variable was whether participants decided to cooperate or compete in the prisoner's dilemma game.

Procedure

Participants joined a Zoom video session on their home computers. A researcher guided the participants through the study. Participants were asked to mute their own audio and video. The Zoom chat feature was used to message information to participants to coordinate gameplay. After providing informed consent, participants worked through a series of counter-balanced scales (i.e., presented in a random order): belief in a just world, ambiguity tolerance, system justification, and the PANAS. Next, they completed a tedious task that involved counting letters on a textbook page (e.g., count all occurrences of the letter "E" on a page). This was intended to distract participants from drawing any connection between the survey content and the game play.

Next, participants were paired at random but not notified about the name of their partner so researchers could assign pairings even in odd group sizes. In the game, participants were told there was additional letter-counting that needed to be completed. Participants were told they had two options: they could assign their partner to do the task, or they could split the task between them. Participants were aware of the payout matrix (Table 1) but unsure what their payout would be as the payout is determined by both partners' choices — not unilaterally. Payouts, or consequences for the game, included zero, one, two, and four pages of letter-counting tasks similar to that which was completed as the distraction task. See Table 1 for the 'payout' structure presented to participants. After deciding, participants completed the letter-counting task per the payout matrix.

Player	**Matrix Result**	**Letter Counting Payout**
Player 1	Split	1 page
Player 2	Split	1 page
Player 1	Split	4 pages
Player 2	Assign	0 pages
Player 1	Assign	0 pages
Player 2	Split	4 pages
Player 1	Assign	2 pages
Player 2	Assign	2 pages

Table 1. Prisoner's Dilemma Payout Matrix as shown to participants

Results

It was hypothesized that people high in aspects of cosmic uncertainty would be more likely to compete (vs. cooperate). Thus, a multiple logistic regression model was conducted to predict people's decision to assign or split from their belief in a just world, ambiguity tolerance, and system justification (Table 2). As this was an initial exploratory study, a liberal measure of statistical significance was accepted (α = .10).

	Odds Ratio	95% Confidence Interval	p-value
BJW	1.25	(0.68, 2.31)	0.46
AT	1.29	(0.54, 3.04)	0.55
SJS	1.42	(0.96, 2.10)	0.07

Table 2. Study 1 Results. Note: BJW = belief in a just world, AT = ambiguity tolerance, SJS = system justification scale

In Study 1, most people chose to cooperate (split the letter-counting work: 69.5%) rather than compete (assign the letter-counting work: 31.5%). Neither ambiguity tolerance nor belief in a just world predicted people's decisions to compete or cooperate (both p > .10). However, initial evidence was found that there is a correlation between a participant's system justification scale score and their decision to compete ($p = 0.07$). People who scored highly in system justification were more likely to assign letter-counting work to their fellow participant (Figure 1). Although only marginally significant by conventional standards, this preliminary finding warrants further investigation.

Study 1 carried several limitations. After analyzing Study 1's results, it was concluded that low internal reliability in the belief in a just world and ambiguity tolerance scales could be driving (null) results. In other words, these individual differences may not have been optimally measured. Additionally, an unmeasured individual difference (i.e., religiosity) was identified as a likely driver of justice beliefs. Like people high in system justification, highly religious people generally share a strong sense of cosmic justice. So, highly religious people may also behave in similarly competitive ways. Yet, no religiosity measure was included in Study 1. Finally, Study 1 measured a dispositional sense of cosmic justice but did not manipulate perceptions of cosmic justice. Study 2 was designed to address these limitations.

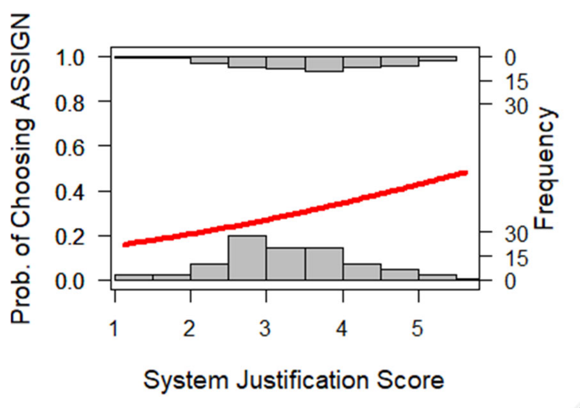


Figure 1. System Justification Scale Results from Study 1

Study 2

Participants

For the second study, students were recruited from introductory psychology courses ($N = 135$) at a large midwestern university for course credit. Participants were 55% Male and 45% Female. The racial background was 63% White, 18% Asian, and 15% Other. Participants ranged from 17 to 35 years old ($M = 19.3$, $SD = 1.8$). Due to time limitations, some participants were unable to answer demographic questions at the end of the study (14 no response to gender and 16 no response to race; percentages reported exclude non-respondents). Participants signed up for times resulting in groups of no less than three and no more than ten.

Materials

Study 2 tasked participants to complete scales for five constructs: Global Belief in a Just World,[21] Ambiguity Tolerance,[22] System Justification,[18] a Religion Index,[23] and a distraction task (completing a Positive Negative Affect scale).[20] Note the changes from Study 1: a revised (global) belief in a just world scale,[21] a revised ambiguity tolerance scale,[22] and the addition of a religion index.[23] The revised scales were included to improve internal reliability (i.e., improve measurement). The religion index was included to explore another possible basis of behavior under uncertain conditions. Internal reliability as measured by Cronbach's alpha was high for belief in a just world ($\alpha = .89$), ambiguity tolerance ($\alpha = .82$), system justification ($\alpha = .84$), and religiosity ($\alpha = .91$).

The letter-counting task was identical to that used in Study 1.

Design

Study 2 aimed to experimentally manipulate cosmic uncertainty. Participants were randomly assigned to either a fair or unfair condition. The fair condition asked participants to write a detailed paragraph about, "a time in your life when you felt the world was very fair to you and that you generally got what you deserved." The unfair condition asked participants to write a detailed paragraph about, "a time in your life when you felt the world was very unfair to you and that you generally did NOT get what you deserved." These paragraphs were designed to manipulate participants' sense of cosmic fairness and order. This design was adapted from a commonly used power manipulation wherein participants write about a time they experienced high or low power.[24] Thus, independent variables in Study 2 included the experimental condition and the four scales measuring beliefs about uncertainty. The dependent variable was whether participants decided to cooperate or compete in the prisoner's dilemma game.

Procedure

As in Study 1, participants were guided by a researcher over Zoom. After providing informed consent, participants worked through a series of counter-balanced scales: belief in a just world, ambiguity tolerance, system justification, a religious index, and a distraction scale. Next, they completed one page of letter-counting as in Study 1 (which introduced participants to the tedious work they would later assign or split). This was followed by a writing task about a fair or unfair experience in the participant's own life. Next, participants went through the same procedure as Study 1's prisoner's dilemma type game: Participants could assign their partner to do additional letter-counting, or they could split the counting between them.

Results

Study 2 results were expected to replicate our Study 1 results in that people high in facets of cosmic uncertainty would be more likely to compete (vs. cooperate). Thus, a multiple logistic regression model was conducted that predicted people's decision to assign or split from their belief in a just world, ambiguity tolerance, system justification, and religiosity (Table 3). The researchers also examined whether the fairness manipulation made participants more or less likely to compete. For these confirmatory analyses, a more conservative threshold of statistical significance was adopted ($\alpha = .05$). Finally, exploratory analyses were conducted to assess whether demographic variables such as age or gender predicted a participant's decision to compete (or cooperate). For these exploratory analyses, a more liberal threshold of statistical significance was used ($\alpha = .10$)

	Odds Ratio	95% Confidence Interval	p-value
BJW	1.82	(1.16, 2.97)	0.01*
AT	1.51	(0.71, 3.33)	0.29
SJS	1.80	(1.22, 2.74)	0.004**
RI	0.93	(0.68, 1.28)	0.67

*Signif. codes: **' 0.01 '*' 0.05*

Table 3. Study 2 Results. Note: BJW = belief in a just world, AT = ambiguity tolerance, SJS = system justification scale, RI = religious index

As in Study 1, most people chose to cooperate (split the letter-counting work: 69.6%) rather than compete (assign the letter-counting work: 31.4%). The association between system justification and a participant's decision to choose to "Assign" ($p < 0.05$) was replicated. Additionally, there was evidence that people with a higher belief in a just world were more likely to assign letter-counting work ($p < 0.05$). That is, people who scored higher in belief in a just world and system justification were more likely to choose the competitive option of "assigning" work to their peer. As in Study 1, no evidence supported the notion that ambiguity tolerance had any meaningful bearing on a participant's decision in the binary prisoner's dilemma game ($p = 0.29$), nor was there evidence to suggest that religiosity plays a role in this competitive decision ($p = 0.67$).

No evidence was found that experimentally manipulating participants' feelings of cosmic justice impacted their competitive decision, nor did it have any interaction with any other scale (Condition p = 0.61; Condition & Belief in a Just World $p = 0.94$; Condition & System Justification $p = 0.94$; Condition & Religiosity $p = 0.52$).

Age was not a significant predictor of a participant's decision to compete or cooperate ($p = 0.63$). However, the data showed quite the opposite for gender ($p < 0.05$). Gender played a significant role in the decision-making process and was correlated with certain personality constructs. When conducting Welch T-tests to examine gender differences between scales, evidence was found that Men ($M = 3.59$) scored higher than women ($M = 3.00$) in belief in a just world ($p < .001$), marginally higher ($M = 3.30$) than women ($M = 3.15$) in ambiguity tolerance ($p = .10$), and higher ($M = 3.93$) than women ($M = 3.01$) in system justification ($p < .001$). Men ($M = 2.69$) and women ($M = 2.71$) did not differ in religiosity ($p = .92$). Further logistic regression analyses were conducted to assess whether belief in a just world and system justification predicted decisions to compete/cooperate controlling for gender. Both belief in a just world ($p = 0.07$) and system justification ($p = 0.08$) were marginal predictors of competitive behavior even controlling for gender.

DISCUSSION

Across two studies, the overwhelming majority of people chose to cooperate in the face of uncertainty. However, this trend was reversed among individuals who had a strong, stable sense of cosmic certainty — that is, people who believed the world was fair and the system was just tended to compete. This could be due to participants' beliefs about how society works. For instance, those that score highly on system justification are more likely to strongly agree with statements such as 'everyone has a fair shot at wealth and happiness' or 'society is set up so that people usually get what they deserve'.[25] If an individual scores highly on system justification, then it is likely that they believe that they deserve any positive outcomes they receive, as good things only happen for good people. If something drastic were to change in the society that currently serves them well, then they would generally be worse off. Those that believe that the system is working in their favor may experience less (if any) guilt that comes with knowing others are less fortunate. It is worth noting that these effects held for both men and women and were not dependent on participants' ages.

Our findings contrasted to prior studies that suggest that uncertainty leads to competitive behavior. In the classroom study that ranked students, under uncertain conditions students tended to resort to selfish, competitive behavior so as to not jeopardize their own position in the class rankings.[3] In our study, most of our participants chose to cooperate. This pattern of results is more consistent with research showing that people tend to cooperate to improve their reputation for future interactions.[12] However, as participants in the current study played only a single round of the game, this study suggests that reputation-building is not the only driver of cooperation. As one participant in our study stated, they saw cooperation as the 'fair' option and 'didn't want the other person to have so much work left.'

Limitations and Future Directions

Our research involved several limitations. Our study's first limitation was that our ambiguity tolerance and belief in a just world scales yielded low internal reliability scores in Study 1. This limitation was addressed in Study 2 in the form of revised belief in a just world[21] and ambiguity tolerance[22] scales with a history of higher internal reliability. Indeed, internal reliability improved for both belief in a just world ($\alpha = 0.69$ in Study 1; $\alpha = 0.89$ in Study 2) and ambiguity tolerance ($\alpha = 0.61$ in Study 1; $\alpha = 0.82$ in Study 2). Future research may consider replicating these findings with the revised scales. Alternatively, as these scales were intended to measure the novel construct of "cosmic justice," this construct may deserve its own scale rather than be measured through these proxy scales.

Another limitation this study ran into is that, due to COVID-19, all research had to be conducted online through platforms such as Qualtrics and Zoom. Individual participants never saw or interacted with each other as they would in a classroom or workplace. Similarly, participants did not know who they would play the competitive behavior game against. Although these decisions likely bolster internal validity, they limit external validity when applying these concepts to 'real-world' cases. In the 'real world,' people are expected to exhibit even more cooperative behavior, as a number of social pressures manifest when individuals are making decisions face-to-face.[26] For instance, a classic study shows that people are more inclined to help others (1) when they empathize with the target and (2) when they have to see the ramifications of their actions.[24]

Though this research is not definitive, it opens the door for future research to further investigate this relationship. More specifically, researchers should work to better isolate and understand a causal relationship between one's system justification score and their projected behavior. Study 2 attempted to experimentally manipulate perceptions of systemic fairness to little success. Writing one paragraph may not be enough to manipulate a participant's belief in cosmic justice. Future studies may use a stronger manipulation. Such manipulations could include longer free-response essay questions about times when participants were treated fairly or unfairly by the "system" or "society at large." Alternatively, researchers could implement fake news articles or videos that portray examples of society's parity or inequality.

Theoretical and Practical Implications

Though decision uncertainty has been studied at length,[1, 3, 9, 10, 15, 27] cosmic uncertainty, as defined in this study, is a nuanced approach to explain human competitive behavior. The findings suggest that in the face of uncertainty, people who have a stable sense of cosmic justice — i.e., those high in system justification and belief in a just world — tend to compete. Further research is needed to determine the bounds of this effect and how task definitions influence people's perceptions of uncertainty in social games.

This study has further societal implications. Though these concepts may appear abstract, there are real decisions people make in the face of uncertainty on a daily basis. Understanding individual differences in beliefs about cosmic fairness could lead to understanding people's decisions to quarantine, wear masks, get vaccinated, or socially distance. In order for the world to be put back together again, it is vital to understand the inner workings of how decisions are made based on underlying individual differences in beliefs.

ACKNOWLEDGEMENTS

I'd like to give a generous thanks to Zach Reese as this manuscript would not have been possible without his excellent guidance and mentoring at every step of the way. I'd also like to thank Dr. Seifert and Dr. Garcia for their help with the overall direction of the project. This research was conducted through Michigan's Social Comparison and Close Relationships lab and I am extremely grateful for all the research assistants' help with data collection. Thank you to my family and friends for supporting me every step of the way.

REFERENCES

1. Torii, T., & Hidaka, S. (2015) Mutual cooperation out of greedy learners: Prisoner's dilemma under uncertainty. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*, 22(3), 341–355. *https://doi.org/10.11225/jcss.22.341*
2. Zuckerman, M. (1975) Belief in a just world and altruistic behavior. *Journal of Personality and Social Psychology*, 31(5), 972–976. *https://doi.org/10.1037/h0076793*
3. Burleigh, T. J., & Meegan, D. V. (2018) Risky prospects and risk aversion tendencies: Does competition in the classroom depend on grading practices and knowledge of peer-status? *Social Psychology of Education: An International Journal*, 21(2), 323–335. *https://doi.org/10.1007/s11218-017-9414-x*
4. Taroni, F., Bozza, S., Biedermann, A., & Aitken, C. (2016) Dismissal of the illusion of uncertainty in the assessment of a likelihood ratio. *Law, Probability and Risk*, 15(1), 1–16. *https://doi.org/10.1093/lpr/mgv008*
5. Braun, P., Cuzzocrea, A., Leung, C. K., & Pazdor, A. G. M. (2018) Item-centric mining of frequent patterns from big uncertain data. *Procedia Computer Science*, 126, 1875–1884. *https://doi.org/10.1016/j.procs.2018.08.075*
6. Baron, D. P. (1970) Price Uncertainty, Utility, and Industry Equilibrium in Pure Competition. *International Economic Review*, 11(3), 463–480. *https://doi.org/10.2307/2525325*
7. Wernerfelt, B., & Karnani, A. (1987) Competitive strategy under uncertainty. *Strategic Management Journal*, 8(2), 187–194. *https://doi.org/10.1002/smj.4250080209*
8. Zinn, J. O. (2009) The sociology of risk and uncertainty: A response to Judith Green's "Is it time for the sociology of health to abandon 'risk'?" *Health, Risk & Society*, 11(6), 509–526. *https://doi.org/10.1080/13698570903329490*
9. Danczak, A., & Lea, A. (2017) The psychology of uncertainty in difficult decisions. *InnovAiT*, 10(8), 466–472. *https://doi.org/10.1177/1755738017709034*
10. Sleesman, D. J., & Conlon, D. E. (2017) Encouraging prosocial decisions: The role of fairness salience and uncertainty. *Journal of Behavioral Decision Making*, 30(2), 502–515. *https://doi.org/10.1002/bdm.1970*

11. Buston, P. M., & Balshine, S. (2007) Cooperating in the face of uncertainty: A consistent framework for understanding the evolution of cooperation. *Behavioural Processes*, 76(2), 152–159. *https://doi.org/10.1016/j.beproc.2007.01.020*

12. Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1996) Cooperation without Reputation: Experimental Evidence from Prisoner's Dilemma Games. *Games and Economic Behavior*, 12(2), 187–218. *https://doi.org/10.1006/game.1996.0013*

13. Hardin, G. (1968) The Tragedy of the Commons. *Science*, 162(3859), 1243–1248. *DOI: 10.1126/science.162.3859.1243*

14. Balafoutas, L., & Sutter, M. (2019) *How Uncertainty and Ambiguity in Tournaments Affect Gender Differences in Competitive Behavior,* 118, 1-13.

15. Vives, M.-L., & Feldman-Hall, O. (2018) Tolerance to ambiguous uncertainty predicts prosocial behavior. *Nature Communications*, 9, 1–9. *https://doi.org/10.1038/s41467-018-04631-9*

16. Rubin, Z., & Peplau, A. (1973) Belief in a Just World and Reactions to Another's Lot: A Study of Participants in the National Draft Lottery1. *Journal of Social Issues*, 29(4), 73–93. *https://doi.org/10.1111/j.1540-4560.1973.tb00104.x*

17. Herman, J. L., Stevens, M. J., Bird, A., Mendenhall, M., & Oddou, G. (2010) The Tolerance for Ambiguity Scale: Towards a more refined measure for international management research. *International Journal of Intercultural Relations*, 34(1), 58–65. *https://doi.org/10.1016/j.ijintrel.2009.09.004*

18. Jost, J. T., & Banaji, M. R. (1994) The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology*, 33(1), 1–27. *https://doi.org/10.1111/j.2044-8309.1994.tb01008.x*

19. Doebeli, M. and Hauert, C. (2005), Models of cooperation based on the Prisoner's Dilemma and the Snowdrift game. *Ecology Letters*, 8: 748-766. *https://doi-org.proxy.lib.umich.edu/10.1111/j.1461-0248.2005.00773.x*

20. Watson, D., Clark, L. A., & Tellegen, A. (1988) Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070. *https://doi.org/10.1037/0022-3514.54.6.1063*

21. Lipkus, I. (1991) The construction and preliminary validation of a global belief in a just world scale and the exploratory analysis of the multidimensional belief in a just world scale. *Personality and Individual Differences*, 12(11), 1171–1178. *https://doi.org/10.1016/0191-8869(91)90081-L*

22. McLain, D. L. (2009) Evidence of the Properties of an Ambiguity Tolerance Measure: The Multiple Stimulus Types Ambiguity Tolerance Scale–II (MSTAT–II). *Psychological Reports*, 105(3), 975–988. *https://doi.org/10.2466/PR0.105.3.975-988*

23. Koenig, H. G., & Büssing, A. (2010) The Duke University Religion Index (DUREL): A Five-Item Measure for Use in Epidemiological Studies. *Religions*, 1(1), 78–85. *https://doi.org/10.3390/rel1010078*

24. Galinsky, A. D., Magee, J. C., Inesi, M. E., & Gruenfeld, D. H. (2006) Power and Perspectives Not Taken. *Psychological Science*, 17(12), 1068–1074. *https://doi.org/10.1111/j.1467-9280.2006.01824.x*

25. Roccato, M. (2014) Measurement properties of the system justification scale: A Rasch analysis. *TPM - Testing, Psychometrics, Methodology in Applied Psychology*, 3, 467–478. *https://doi.org/10.4473/TPM21.4.7*

26. Toi, M., & Batson, C. D. (1982) More evidence that empathy is a source of altruistic motivation. *Journal of Personality and Social Psychology*, 43(2), 281–292. *https://doi.org/10.1037/0022-3514.43.2.281*

27. Schulze, C., van Ravenzwaaij, D., & Newell, B. R. (2015) Of matchers and maximizers: How competition shapes choice under risk and uncertainty. *Cognitive Psychology*, 78, 78–98. *https://doi.org/10.1016/j.cogpsych.2015.03.002*

ABOUT STUDENT AUTHOR

Duncan Drewry was an undergraduate student at the University of Michigan at Ann Arbor and graduated with a B.A. in Cognitive Science with distinction and high honors. While at Michigan he also majored in Spanish and was a member of the Social Comparisons and Close Relationships lab working on Psychology research.

PRESS SUMMARY

How do our underlying beliefs about fairness, justice, and ambiguity influence our behavior? Does it lead us to cooperate or compete? Little research has examined how individual differences between people moderate cooperative and competitive behavior in the face of uncertainty. This study proposes that a stable sense of justice may, ironically, lead to more competitive behavior. Specifically, people who score highly in belief in a just world, system justification, and religiosity may be more inclined to compete rather than cooperate because they see themselves as good and believe they deserve positive outcomes. Our results show that people tended to cooperate, but those who believe the current social system is fair and that the world is just were more likely to compete. In other words, people with a strong sense of cosmic justice would be more likely to exhibit competitive behavior under uncertain conditions.

Print Edition ISSN 1536-4585
Online Edition ISSN 2375-8732