
Incorporating Adaptive Human Behavior into Epidemiological Models using Equation Learning

Austin T. Barton¹ Jordan Klein² Kevin B. Flores^{3*}

¹Georgia Institute of Technology, Atlanta, GA, USA

²Appalachian State University, Boone, NC, USA

³Department of Mathematics, North Carolina State University, Raleigh, NC, USA

abarton40@gatech.edu kleinja@appstate.edu kbflores@ncsu.edu

*Corresponding author.

Abstract

Mathematical models have proven to be valuable tools for forecasting and studying epidemics. Covasim is an open-source agent-based model (ABM) developed to simulate the transmission of COVID-19. Human behaviors, such as compliance to masking, have been shown to depend on the state of the epidemic and can significantly impact disease spread. While Covasim's base model does not include adaptive elements, we extended the Covasim model to incorporate adaptive masking behavior to investigate its effect on the system's behavior, complexity, and our ability to obtain an ordinary differential equation (ODE) approximation from it. Using an existing compartmental model, we processed the data generated from this extended ABM through Biologically-Informed Neural Networks (BINNs) and sparse regression methods to obtain an ODE approximation. The extended ABM and equation learning pipeline we developed is open-source to provide a quantitative framework for incorporating adaptive behaviors into forecasting future epidemics and other similar computational models.

1 Introduction

Epidemics present a multi-faceted challenge characterized by highly complicated and intricate interactions between individuals and a disease within a population. These interactions and disease dynamics form an epidemiological system. In addition to such a system often being highly complicated, the population that the epidemic affects may be composed of individuals that act adaptively (or maladaptively) to the state of the disease and the epidemic as a whole. Human behaviors, such as masking, that occur in response to an epidemic, such as COVID-19, can significantly affect the way the disease spreads [1, 2]. In this work, the term "masking" refers to the act of wearing a facial covering for the purpose of preventing the spread of disease. This additional aspect of complexity may significantly increase the difficulty of accurately modeling disease spread and necessitates novel approaches that capture both the individual-level interactions and the population-wide dynamics of the disease spread that result from such behaviors.

Traditional epidemiological models, while providing valuable insights, often struggle to account for individual-level heterogeneity and dynamic adaptation such as masking, quarantining, or working remotely in response to disease spread. These models typically rely on compartmental models that represent the population in aggregate states, such as susceptible, infected, and recovered (SIR) [3]. While these models offer a simplified view and can be computationally efficient, they often lack

the necessary granularity to capture the nuances of individual responses and their impact on the overall epidemic trajectory. Agent-based models (ABMs), and other computational models, offer unique advantages over traditional modeling approaches. ABMs excel at capturing complex system dynamics driven by individual-level interactions, heterogeneity, and emergent behavior, providing valuable insights into real-world phenomena [4, 5]. These benefits extend to various fields, including biology [6], public health [7], and social sciences [8]. ABMs' flexibility also allow us to study and gain deeper insights into more complex phenomena without losing its capabilities of accurately modeling a system. Specifically, agent behaviors within the model can be programmed to adapt to the system it resides in based on certain factors and information [9]. However, computational models also face limitations. Despite their ability to represent intricate systems, they can be computationally expensive, particularly with increasing complexity [10, 11]. Additionally, model validation and ensuring biological plausibility can be challenging tasks [12, 6, 10]. Furthermore, analyzing outputs from complex models often requires advanced statistical techniques and careful consideration of variance and sensitivity [11].

There are many reasons one would prefer to use a differential equation model due to its computational efficiency over most ABMs and our ability to quantitatively analyze them for equilibria, stability, sensitivity to certain parameters, etc. However, deriving an ODE that is capable of accurately modeling a system with adaptive behaviors is difficult and often intractable. Nardini et al., in their 2021 paper titled: "Learning differential equation models from stochastic agent-based models", demonstrate and analyze numerous methods of deriving a differential equation model from an ABM [12]. Thus, allowing one to leverage the flexible and robust framework of ABMs to simulate complex dynamic behavior and use them to obtain mathematical models. In their paper, they discuss two common methods of predicting emergent behaviors of ABMs: Monte Carlo methods and coarse-grained modeling. However, Monte Carlo methods require extensive simulation of many ABMs which may not be feasible and coarse-grained modeling often requires simplifying assumptions to approximate its dynamics with a more tractable DE model. The most commonly used differential equation (DE) model approximations for ABMs are mean-field models [13, 14, 15, 16]. Mean-field DE models are often simple to solve, so they provide an advantageous alternative to extensive simulation of the ABM. However, mean-field models require the *mean-field assumption*, which assumes that the actions of agents are independent of one another. However, the adaptive behaviors we are studying occur in response to the state of a system and may have a significant effect on the dynamics of the system, which affects the behavior of the agents. Therefore, such an assumption can be erroneous when studying systems with adaptive behaviors which can negatively affect the derived DE model greatly. Nardini et al. propose a novel approach for inferring DE models directly from stochastic (ABM) simulations [12]. This method leverages equation learning, overcoming limitations of traditional approaches like Monte Carlo simulations or coarse-grained models. The authors utilize equation learning to avoid extensive simulations while still being able to capture the full complexity of ABM dynamics. Their approach successfully infers DEs from two example ABMs, demonstrating accurate predictions of ABM dynamics and improved computational efficiency compared to classical methods [12].

Such methods are very useful in our context, but a major obstacle still exists when attempting to model a system with added complexities such as adaptive behavior. The dynamics of certain parameters may be drastically affected, making their underlying form unknown to the modeler. This implies that one needs an estimation of certain parameters (sometimes referred to as mechanistic models) before being able to effectively infer an equation on them. The 2020 paper, "Biologically-informed neural networks guide mechanistic modeling from sparse experimental data", by Lagergren et al. explores a novel approach to overcoming such limitations for learning mechanistic models (i.e. equations of parameters) [17]. These traditional methods often require prior knowledge of the underlying mechanisms, which can be limiting in various research fields. They introduce Biologically-Informed Neural Networks (BINNs), which combine parameter approximating neural networks with physics-informed neural networks (PINNs). PINNs are adept at approximating solutions to partial differential equations (PDEs), allowing them to provide good approximate solutions in order to learn the parameters of interest while minimizing *a priori* assumptions (further details in 2.4). We use the framework laid out by Lagergren et al. to infer nonlinear parameters as closed equations in order to obtain a mathematical model that is capable of simulating a system with adaptive elements. This particular approach of inferencing mathematical models from stochastic agent-based models is unique, and although our data is not sparse or experimental, we provide insight into emerging techniques in equation learning while utilizing ABMs as a tool for simulating a system with adaptive

elements. Xin Li's thesis titled, "Artificial Intelligence Models to Improve Organ Allocation and Public Health Decisions", illustrates a similar process where they use BINNs and sparse regression in order to obtain an ODE approximation of Covasim without adaptive behaviors [?]. Our work expands on this by including adaptive behaviors and creating an accessible approach to modeling systems with adaptive behaviors.

1.1 The Computational Pipeline

We developed an end-to-end computational pipeline for augmenting ABMs with adaptive behaviors and processing the data through an equation learning pipeline to obtain an ODE approximate. In our work, we leverage a comprehensive ABM of COVID-19 called Covasim ([18]) and augment it with a masking behavior to simulate the dynamics of a viral epidemic with adaptive elements. For our equation learning pipeline, we then utilize BINNs for estimating nonlinear parameters that abide by the dynamics of our ABM and apply sparse regression to obtain an effective ODE approximation. Combining these approaches, we demonstrate that it is possible to extract a set of governing equations that capture the essence of the system's behavior. This approach offers several advantages, including:

- Predictive power: The equations can be used to make predictions about the future course of the epidemic and assess the potential impact of different control interventions.
- Reduced computational complexity: The resulting set of equations can be significantly more efficient to solve than running an ABM, making it possible to perform simulations capable of meaningful extrapolation and analyze scenarios more quickly.
- Increased interpretability: The explicit equations provide a clearer understanding of the relationships between key variables and the underlying mechanisms driving the epidemic's progression and its various defining parameters, such as contact rate.
- Quantitative Analyses: DE approximations can be further analyzed using a variety of quantitative methods such as sensitivity analysis, studying equilibria, etc.

Combining the strengths of agent-based modeling and equation learning holds significant promise for advancing many different fields similar to epidemiology. This approach is particularly powerful when considering the adaptive behaviors of individuals within the population, as it allows data-driven inferencing to obtain a more accurate representation of the complex interactions that shape the course of an epidemic. By offering a more realistic and tractable approach to disease modeling, this framework has the potential to improve our understanding of epidemic dynamics and contribute to the development of more effective control strategies. In the following work, we showcase an effective end-to-end computational pipeline (also shown in Figure 1) that consists of:

- Selecting ABMs for simulating data.
- Selecting a base mathematical model and identifying potentially nonlinear parameters that need to be estimated.
- Implementing adaptive behavior and augmenting the ABM.
- Utilizing existing frameworks for parameter estimation such as BINNs (Section 2.4).
- Identifying and addressing necessary modifications to the base mathematical model.
- Implementing equation learning techniques for obtaining an ODE approximation of the augmented ABM.

2 Modelling and Equation Learning

2.1 Our Agent-Based Model: Covasim

Covasim, an agent-based simulator created by the Institute for Disease modeling (IDM), offers a valuable tool for simulating and studying disease dynamics [18]. Covasim's features are easily extended and adapted by users to tailor the simulation to specific diseases and scenarios. In the paper by Kerr et al. titled, "Covasim: An Agent-Based Model of COVID-19 Dynamics and Interventions" [18], the simulations provided as examples use a variety of features to model agent interactions and disease dynamics. Agents exist in different network layers including school, workplace, household,

Computational Pipeline

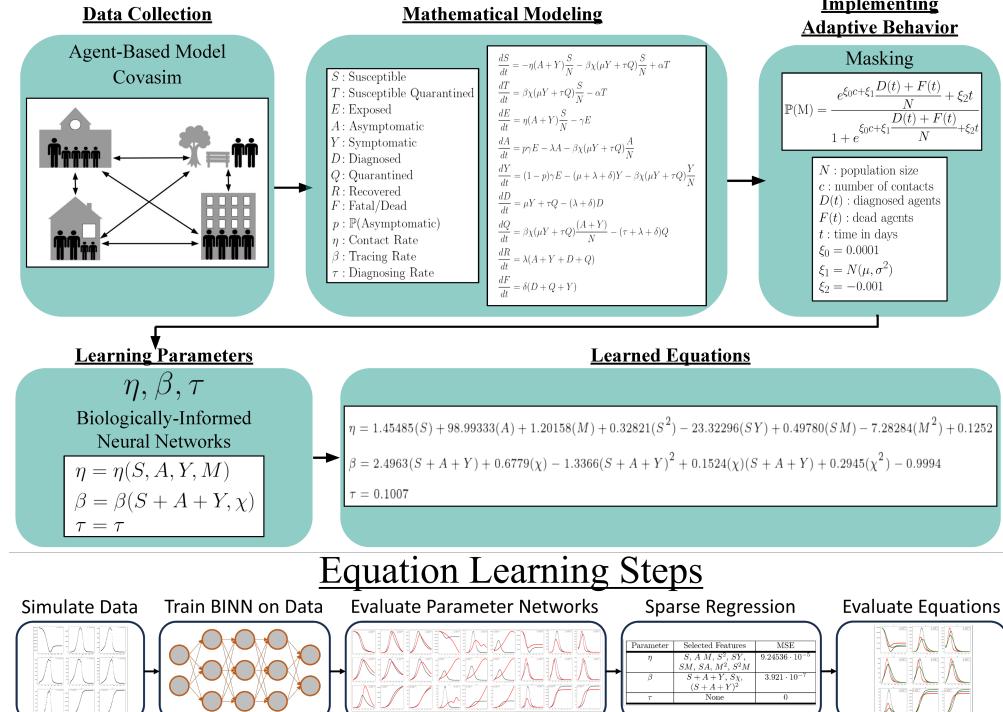


Figure 1: Schematic of the computational and modeling pipeline (top) along with the equation learning data processing steps (bottom). We start with a base computational model and mathematical model, augment each of them by including adaptive behavior, then perform parameter estimation and equation learning to obtain a closed approximate system of differential equations.

and community each with their own sets of values for probability of infection, tracing, quarantining, etc. Despite how comprehensive and capable Covasim is at simulating the disease-spread and population dynamics, the existing model lacks a key element: adaptive behavior of agents.

2.2 Incorporating Adaptive Behavior: Masking

We define an adaptive behavior of an agent to be any action that is dependent on the state of the system it resides in. Typically, this behavior will also have an effect on the dynamics of the system, creating a feedback loop between the system and the agents' behaviors. Since adaptive behavior makes actions of agents dependent on the actions of other agents, it directly violates the mean-field assumption making coarse-grained modeling far more difficult [12].

In real-world disease outbreaks, individual behavior dynamically adapts based on the perceived threat and evolving conditions of the epidemic. This feedback loop between individual actions and the overall system dynamics significantly impacts disease spread and control strategies [19]. We address this by introducing adaptive masking behavior into the Covasim framework. Masking significantly affects disease transmission and is highly dependent on the perceived severity and prevalence of the epidemic [20]. Thus, incorporating this adaptive behavior is essential for achieving a more realistic and accurate model of disease spread.

We introduce a probabilistic masking model based on the diagnosed (D), fatalities (F), and time (t) variables. Intuitively, the probability of masking is expected to be:

- Increasing with D and F : As the number of diagnosed cases and fatalities increase, the perceived risk of infection rises, thereby motivating individuals to adopt masking behavior.
- Decreasing with t : Over time, complacency may set in, leading to a decrease in masking behavior even as the epidemic persists.

We use the logit function with D , F , and t as features to model this adaptive behavior. Our probability of masking for each agent is:

$$p_M = \mathbb{P}(\text{masking}|D, F, t) = \frac{e^{\xi_0 + \xi_1 \frac{D(t) + F(t)}{N} + \xi_2 t}}{1 + e^{\xi_0 + \xi_1 \frac{D(t) + F(t)}{N} + \xi_2 t}} \quad (1)$$

where $D(t) :=$ Number of diagnosed agents at time t , $F(t) :=$ Cumulative number of agents that died from infection at time t , $t \in \mathbb{N}$ and $t :=$ time in days of the simulation (each "day" of the simulation is an iteration), and ξ_0, ξ_1, ξ_2 are hyperparameters summarized as follows:

- $\xi_0 \in \mathbb{R}^+ \cup \{0\}$: Represents the initial proportion of the population masking.
- $\xi_1 \sim \mathcal{N}(\omega, \sigma)$: controls the sensitivity of masking behavior to the epidemic's severity, where $\omega \in \mathbb{R}$ is the mean and $\sigma \in \mathbb{R}^+ \cup \{0\}$ is the standard deviation.
- $\xi_2 \in \mathbb{R}^- \cup \{0\}$: dictates the rate of decrease in masking behavior over time.

Each agent then has a corresponding binary random variable $\mathcal{M}(t)$ where

$$\mathcal{M}(t) = \begin{cases} 1 & \text{If masking at time } t. \\ 0 & \text{If not masking at time } t. \end{cases} \quad (2)$$

And has probability mass function

$$f_M(k; p_M) = \begin{cases} p_M & \text{If } k = 1. \\ 1 - p_M & \text{If } k = 0. \end{cases} \quad (3)$$

When $\mathcal{M} = 1$ at time step t within the simulation, for any given agent, we reduce the effective infection rate across each network layer of the corresponding agent within Covasim to model the effect masking has on reducing the spread of disease [20]. The specific reduction is referred to as the "masking effect" and is a real number in the interval $[0, 1]$ that interacts with the probability of transmission in each network layer of the ABM.

We chose ξ_1 to be sampled from a normal distribution to introduce more stochasticity into the behavior of masking. Intuitively, it is difficult to predict exactly how likely people are to mask in response to rising diagnoses and fatality statistics. Because people can behave irrationally and nondeterministically, adjusting the probability of masking randomly can potentially model this behavior more accurately. Although our methods of modeling masking are effective at changing the dynamics of the ABM, our work doesn't provide a comprehensive guide on accurately modeling this behavior. This limitation of our work is further discussed in Section 4.1.

2.3 Mathematical Model

We use a compartmental model consisting of 9 different states to mathematically model the spread of disease in the simulation and utilize equation learning methods in order to infer certain parameters of our system of equations (similar to Li, Xin [?]). The base model includes three parameters: contact rate (η), tracing rate (β), and quarantine diagnosis rate (τ), that we suspect to be nonlinear functions which we wish to estimate via neural networks (details in Section 2.4). Other parameters are fixed as constants and are directly inferred from our knowledge of hyperparameter values of the Covasim simulation (Table 3).

We hypothesized that introducing such adaptive behavior(s) will likely increase the nonlinearity and complexity of the parameters. In Section 3, we demonstrate and confirm this. Additionally, our implementation of masking directly interacts with and decreases the probability of transmission. Therefore, it is very likely that in the ABM with adaptive behavior, the dynamics of the contact rate (η) are affected and partly determined by the proportion of the population masking, M , in addition to S , A , and Y as assumed in the original model. Because of this, we augment the existing compartmental model to include a non-mutually exclusive compartment representing the proportion of the population masking, M , as shown in Figure 3. Using sophisticated techniques in order to accurately learn parameters such as contact rate, tracing rate, and quarantined rate, is vital to the accuracy and interpretability of the learned system of equations. We used Biologically-Informed Neural Networks (BINNs) in order to do this. For data generated with masking as an adaptive behavior, we include the proportion of population masking as an input into the contact rate parameter.

Compartment	Description	Parameter	Description
S	Susceptible	$\eta(S, A, Y)$	Contact Rate
T	Quarantined and susceptible	$\beta(S + A + Y, \chi)$	Tracing Rate
E	Exposed	$\tau(A, Y)$	Quarantined Diagnosis Rate
A	Asymptomatic and infected	p	Asymptomatic Probability
Y	Symptomatic and infected	γ	Infection Rate
D	Diagnosed and infected	α	Unquarantining Rate
Q	Quarantined and infected	λ	Recovery Rate
R	Recovered	δ	Fatality Rate
F	Fatality	χ	Tracing Interaction (Piecewise)

(a) Compartments of the STEAYDQRF model we use to model Covasim. (b) Parameters of the baseline ODE system.

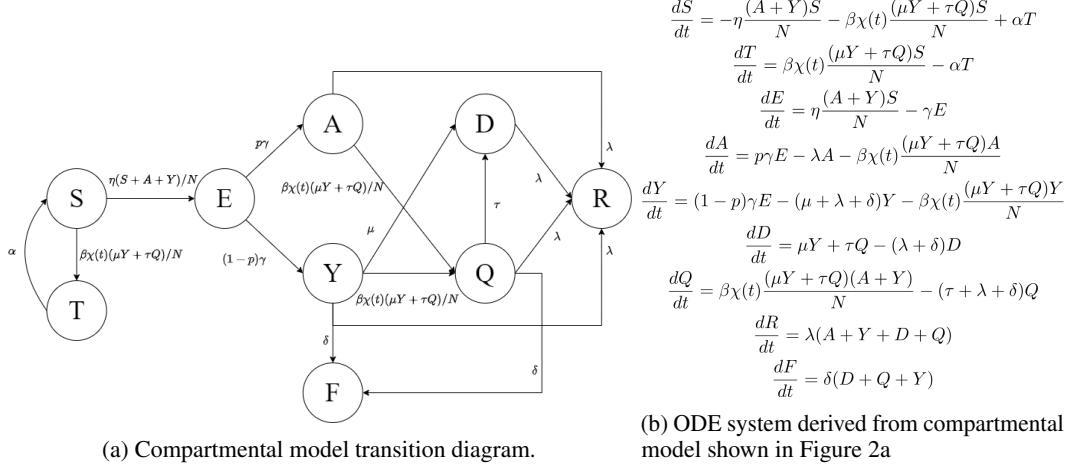


Figure 2: Summary of the base mathematical model we use to model Covasim.

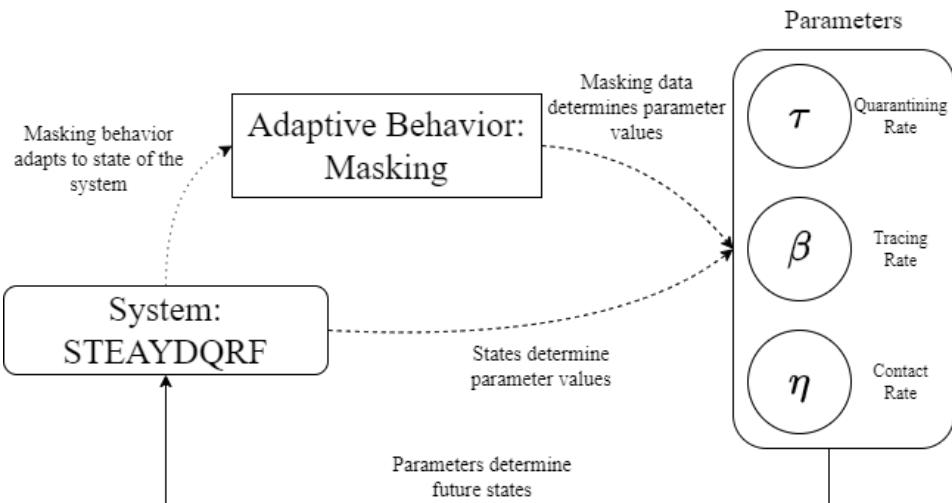


Figure 3: Flowchart diagram illustrating the feedback loop between the states of the system, the adaptive behavior (masking), and the parameters of our model.

2.4 Biologically-Informed Neural Networks for Parameter Estimation

Biologically-Informed Neural Networks (BINNs) attempt to tackle the *model specification problem* by using neural networks as surrogate models to approximate unknown and possibly nonlinear parameters in order to minimize *a priori* assumptions about the form of the differential equation(s) as well as expand the number of possible solutions to such a model [17]. Generally, these parameters may be linear, nonlinear, or constant functions. This makes the space of possible functions too large to accurately solve for without making large assumptions using methods of sparse regression. Hence, using a universal function approximator (such as an MLP [21]) allows us to learn these nonlinear components that adhere to a governing system of differential equations and then use equation learning with feature selection to infer parsimonious equations and underlying relationships *a posteriori*. It is up to the modeler to decide what parameters are worth approximating via parameter networks, which others can be estimated using other methods, or are already known.

The architecture of the BINN is primarily composed of 2 parts:

- 1) The informed-neural network which approximates the solutions to the governing dynamical system (DE). This decides what is inputted into the BINN, what the governing dynamical system is, and how part of loss function is constructed.
- 2) Parameter networks approximating the parameters we wish to estimate. Various parameters may or may not need estimation via parameter networks. These also determine how part of the loss function is constructed.

We construct our BINN by following the blueprint discussed and shown in detail in Lagergren et al.'s paper [17]. We approximate η , β , and τ with parameter networks $\hat{\eta}_{NN}$, $\hat{\beta}_{NN}$, and $\hat{\tau}_{NN}$, respectively. The BINN loss function for the BINN is composed of 3 main components:

- 1) Loss between the predicted solutions and observed data.
- 2) Loss between the derivatives (obtained through backpropagation) of the solutions and the right-hand side value with the parameter networks plugged in.
- 3) Penalty of parameter values outside their expected ranges.

Expert knowledge and intuition are used in order to determine what lies in the domain of the parameter and thus, what is inputted into the parameter network. We assumed the contact rate to be a function of S , A , Y , tracing rate to be a function of $S + A + Y$, $\chi(t)$, and quarantined diagnosis rate to be a function of A , Y . We restrict the parameters to be within the following intervals, $\eta \in [0, 0.4]$, $\beta \in [0, 0.3]$, and $\tau \in [0, 0.6]$ by penalizing values of such parameters outside of their respective codomains in the loss function (details in [17]).

Because, as we will later show in Section 3, incorporating certain adaptive behaviors such as masking into our model drastically changes the dynamics, we include observed data of masking into the contact rate parameter in order to obtain more accurate dynamics from our mathematical model.

2.5 Equation Learning: Sparse Regression

Parameters obtained via trained BINNs are neural networks which approximate the unknown true parameters of the system which we refer to as parameter networks. These parameter networks are non-interpretable function approximations. Although they can provide good approximations of non-linear functions, they do not provide much mathematical insight into the dynamics of the parameter it is approximating. Of course, we could explicitly express these parameter networks as a composition of functions and affine transformations, but they are not very interpretable or useful for analyzing and studying in this way.

We can use supervised learning techniques to approximate these parameter networks in terms of the inputs, often referred to as the library of potential terms for inclusion in the inferred differential equation model. This is described in detail in Nardini et al.'s 2021 paper where it is referred to as the third step of the equation learning pipeline [12]. Specifying this library of terms to act as the basis of this learned equation is done using expert knowledge and intuition of the parameter we wish to estimate. Since the parameter network is a function of specific inputs, we restrict our library to consist of transformations of those inputs. We assume that the parameter networks can be sufficiently

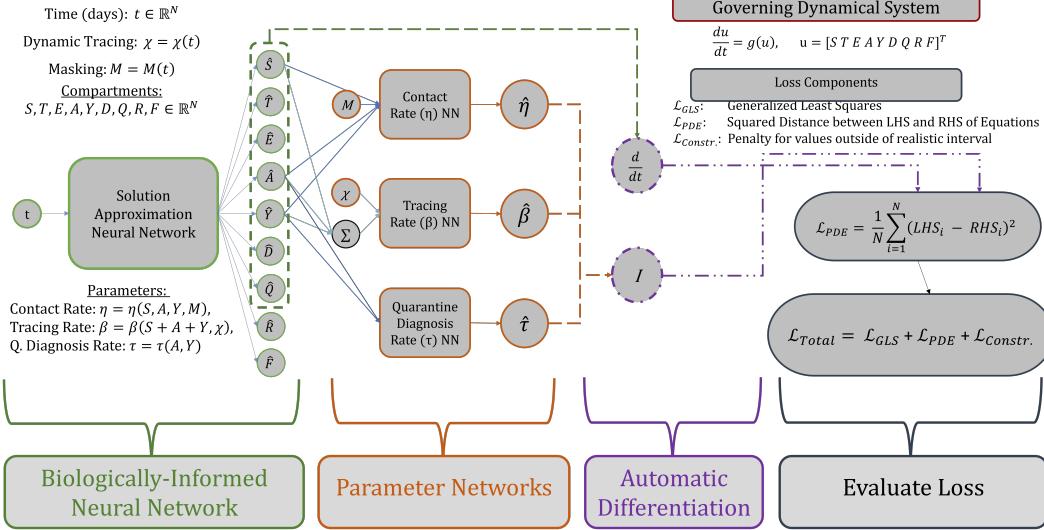
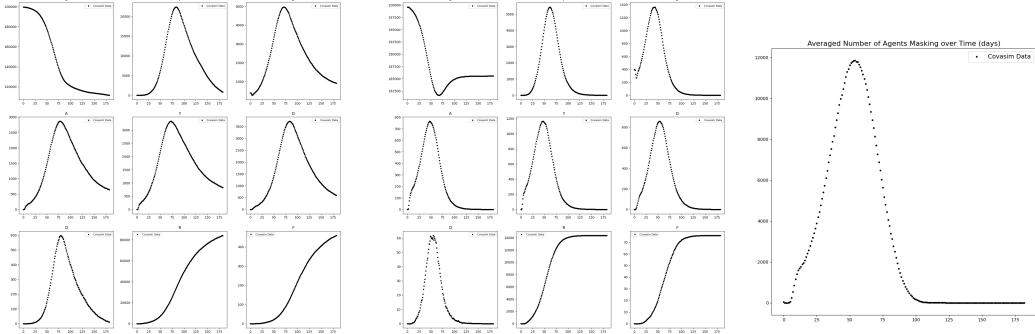


Figure 4: Schematic of our Biologically-Informed Neural Network model for our epidemiological model with masking data as input into the contact rate parameter (η). See Lagergren et al.’s 2020 paper for further details on this architecture [17].

estimated as a polynomial of their inputs restricted to a degree of 4. This choice is somewhat arbitrary and ideally, we would use insights of experts to guide the construction of this library of term. In order to ensure a parsimonious learned equation, we use sparse regression methods. Specifically, we use LASSO regression due to the sparsifying effect of L^1 regularization and its feature selection capabilities [22].



(a) Averaged data generated from Covasim ABM with no masking behavior. The nine figures are the number of agents in each compartment of the STEAYDQRF model over time. This data was averaged over 16 simulations to reduce noise in data.

(b) Averaged data generated from Co-vasim ABM with masking behavior. The nine figures are the number of simulations in each compartment of the STEAYDQRF model over time.

(c) Proportion of population masking over time averaged over 16 agents in each compartment of the STEAYDQRF model over time.

Figure 5: Comparison between data generated by Covasim without and with adaptive masking behavior.

3 Results

In the following section, we give an overview of our data simulation and equation learning pipeline along with an analysis, interpretation, and discussion. Because the ABM is stochastic, the data from each simulation has noise that can make training BINNs and estimating dynamics such as contact rate

more difficult. Parallel computing and Covasim's existing model provide a feasible way to compute multiple ABM simulations at the same time provided that the necessary hardware, such as a graphics processing unit (GPU), is available. Since such hardware is quite common and was available for us, we simulated a batch of 16 Covasim simulations and averaged the data for each compartment and masking for each scenario to reduce noise in the data. In the following results, comparisons are made between two scenarios: no masking behavior included and masking behavior included. Importantly, the only difference between the two scenarios simulated is the adaptive masking behavior, keeping all other hyperparameters of the ABM fixed.

3.1 Effects of Adaptive Behavior on ABM

We first generated data from two different scenarios, no masking and masking, on a population of 200,000 agents in order to first verify and study the effects that masking has on our ABM. Adding adaptive behavior into the agent-based model significantly impacted the dynamics of the system. The model exhibited different behavior in each compartment, as seen in Figure 5 when compared to 5a. Most notably, the susceptible compartment went from having a simple monotonic decreasing "S" curve throughout the simulation to a much different shape with a shift from a negative slope to a positive slope to a near zero slope. Additionally, the final number of agents that were exposed and infected by the disease after the entire simulation in the ABM with masking was drastically less than the simulation with no masking behavior. Specifically, Compartments such as F went from having approximately 1200 agents in it (for no masking) by the end versus just under 200 in the same simulation with masking behavior added. The structure of the curves in each compartment also changed. Lastly, all of the compartments in the scenario with masking visually have much steeper curves.

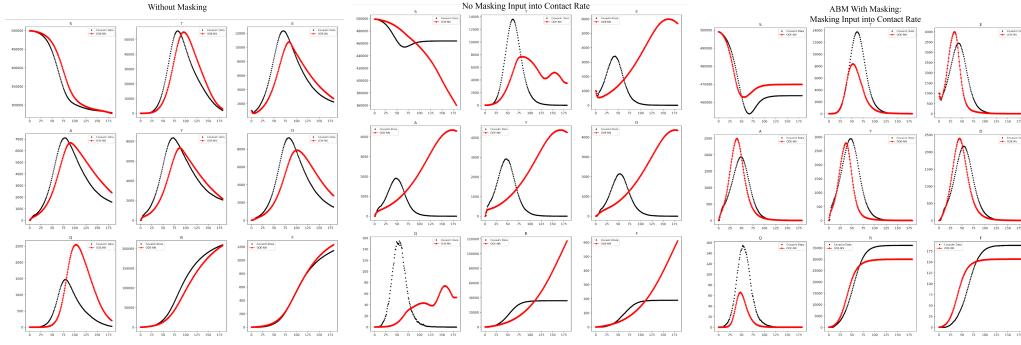
This demonstrates that the adaptive behavior we introduced has a significant effect on the dynamics of the ABM. Moreover, the masking behavior itself also changes in response to the state of the epidemic, as seen in Figure 5. Masking quickly increases from approximately 0 to 2000 (0% to 0.4% of the total population) in response to increased diagnoses and fatalities of the epidemic. As the proportion of the population masking increases, disease spread slows. As the number of diagnosed, $D(t)$, decreases and the cumulative fatalities, $F(t)$, plateau due to agents masking (which directly decreases the probability of transmission in the ABM), masking rates decline over time towards 0. This feedback loop (shown in Figure 3) between the adaptive behavior and the system which the agents reside in is precisely what we are interested in studying due to the complexity to the system it adds.

After analyzing and verifying the effects of our masking behavior on Covasim, we generated data for each scenario (no masking and masking) with 500,000 agents to process through our equation learning pipeline. Details of the hyperparameters for our generated data are outlined in the Appendix A.

3.2 Approximate System of ODEs with Neural Networks

We performed our first step of equation learning by employing BINNs to obtain effective approximations of our parameters of interest via parameter networks. These trained parameter networks effectively capture the dynamic behavior of the parameters and provide solutions that are capable of accurately modeling Covasim with no adaptive behaviors, as seen in Figure 6. However, after introducing masking into Covasim, our implementation of BINNs was incapable of learning parameter networks that sufficiently model the dynamics of the system, as seen in Figure 6. To address this we included masking data as a feature of the contact rate parameter network, $\hat{\eta}_{NN}$. This simple modification to our BINNs provided dramatic improvements in the approximate ODE's ability to model our ABM with adaptive masking behavior, as seen in Figure 6.

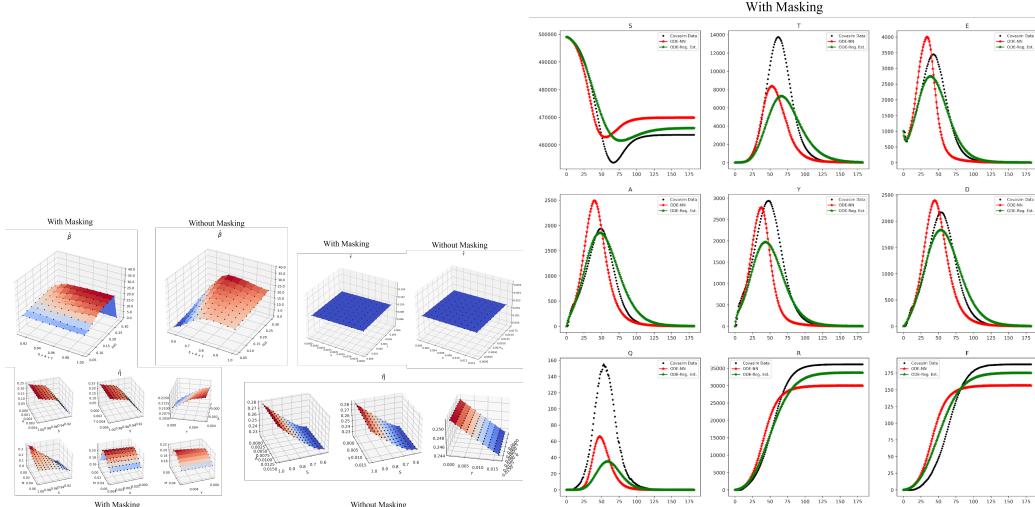
We added masking data as an additional feature to our contact rate parameter because we hypothesized that without any knowledge of masking within the population, the base mathematical model was not able to capture the dynamics of the spread of disease throughout the simulation (as seen in Figure 6). We believed it was likely that certain parameters of the model are directly affected by masking. Specifically, since our implementation of masking affects the probability of transmission for agents across each of their network layers, we believed that masking would affect the contact rate parameter the most. Therefore, we take observed masking data throughout the simulation at each



(a) Comparison between Covasim data without adaptive masking behavior and the solutions of an ODE and the solutions of an ODE approximation with parameter networks. (b) Comparison between observed data with adaptive masking behavior and the solutions of an ODE approximation with parameter networks. (c) Comparison between observed data without masking input into contact rate networks with masking data input into any of the parameters of the model. putted into the contact rate parameter network $\hat{\eta}_{NN}$.

Figure 6: Comparison between observed data (black) and solutions of ODE approximation (red) with parameter networks $\hat{\eta}_{NN}$, $\hat{\beta}_{NN}$, and $\hat{\tau}_{NN}$.

timestep and feed it as an input into the contact rate parameter network, now making our contact rate a function of S , A , Y , and M . This incorporation of adaptive behavior also led to a drastic change in the complexity and nonlinearity of the model’s parameters (discussed in 3.3).



(a) Comparison of surfaces of our parameter networks when trained on Covasim without masking (and no masking behavior, solutions to ODE system with masking input into contact rate) and Covasim with masking input into contact rate. (b) Comparison between observed Covasim data with inferred equations approximating parameters.

Figure 7: Obtained parameter networks with masking adaptive behavior and their corresponding solutions to our ODE system.

3.3 Interpreting Parameters and Sparse Regression for Equation Inferencing

After training the BINN, we have effectively obtained an approximate ODE where the parameters η , β , and τ of the original system of equations are approximated by surrogate models $\hat{\eta}_{NN}$, $\hat{\beta}_{NN}$, $\hat{\tau}_{NN}$, each of which are the trained parameter networks from the trained BINN. In

Parameter	Features with Non-zero Coefficients	MSE
η	$S, A M, S^2, SY, SM, SA, M^2, S^2M$	$9.24536 \cdot 10^{-5}$
β	$S + A + Y, S\chi, (S + A + Y)^2$	$3.921 \cdot 10^{-7}$
τ	None	0

Table 2: Selected features after performing equation learning with LASSO regression on parameter networks. Mean Squared Error (MSE) is measured as the raw MSE between the parameter network and the inference equation.

the previous section (Section 3.2) we demonstrate that the trained parameter networks provide estimates that sufficiently adhere to the dynamics of the system (Figure 6). We could express the parameter networks explicitly with the learned weights, however, there is very little interpretability of such an expression and it is very tedious. Visualizing the surfaces of the parameter networks over their respective domains can provide valuable insight into the dynamics of the parameters. Figure 7a shows such visualizations and compares the learned parameters between the model without masking behavior and with masking. We can see the contact rate dynamics are drastically different between the two scenarios. In addition to the contact rate being a function of masking data in the masking scenario, its relationship between common inputs, such as S , and Y , are significantly different. Additionally, surfaces plotted with M included show the parameter’s dynamics are heavily dependent on the values of M . Furthermore, the tracing rate was also affected by this added adaptive behavior, as shown in Figure 7a. The quarantined diagnosis rate is a special case where the parameter network consistently outputted a constant value for its approximated parameter in both cases. Even though the dynamics of the parameter are the same, we can see that in the scenario with masking, the outputted value is 0.1 whereas in the no masking scenario it is 0.05, indicating that although the dynamics of the parameter didn’t change, masking still had an effect on the values of the parameter.

3.4 Approximate System of ODEs with Parsimonious Equations

After obtaining an approximate ODE with parameter networks, verifying its ability to model the ABM with masking behavior, and analyzing the surfaces of the parameter networks, we undergo sparse regression on the parameter networks to obtain parsimonious equations for each parameter. The library of potential terms we used for each parameter is a polynomial of its inputs up to degree 4. LASSO regression was performed over the domain of each respective parameter with a 80% and 20% split for training and test data, respectively. LASSO regression was conducted over a range of penalty values of the L^1 regularizer to find the optimal weight for the L^1 penalty. Table 2 shows the features with non-zero coefficients after this procedure for each parameter and its corresponding mean-squared error (MSE). The MSE values are quite small, however, this metric for loss alone is not sufficient for evaluating how well these equations estimate the parameter networks. In order to verify that the obtained equations are sufficiently close estimates of the parameter networks, we compare the solutions of the corresponding ODE approximation between the observed data and the solutions with parameter networks plugged into the ODE. The results of this procedure are shown in Figure 7b. As we can see, there is a significant difference between the solutions of the ODE with equations versus the ODE with parameter networks. However, the obtained ODE approximation performs similarly and is capable of capturing the ABM’s dynamics in most compartments.

4 Discussion and Conclusion

This study explored the potential of merging agent-based modeling (ABM) and equation learning to study complex systems with adaptive behavior. Utilizing Covasim, we simulated an epidemic with adaptive masking behavior, demonstrating the significant impact of individual-level behavior on disease dynamics [4]. Biologically-Informed Neural Networks (BINNs) proved effective for parameter estimation, guiding equation learning to infer an effective approximate system of differential equations [17].

Through equation learning, we successfully extracted a closed system of ordinary differential equations (ODEs) from the trained neural network, offering a simplified and interpretable representation of the ABM’s essential behavior [12]. Notably, the original ODE system without masking became

insufficient to capture the dynamics when adaptive behavior was introduced. However, augmenting the model with masking data as a feature of the contact rate parameter resolved this issue.

The increased complexity and nonlinearity introduced by the adaptive behavior were evident in the compartment plots seen in Figure 7b and the surfaces of the parameter networks in Figure 7a. This highlights the challenges associated with modeling such systems. Nevertheless, applying equation learning techniques yielded a parsimonious closed-form system of ODEs that approximated the complex parameter networks. This system provides an interpretable and computationally efficient representation of the ABM, capturing the essential dynamics with a reduced number of terms.

This research contributes to the growing field of equation learning and its application to complex adaptive systems. It demonstrates that combining ABMs and equation learning can be a valuable approach to understand the dynamics of such systems and develop accurate and interpretable models.

4.1 Limitations

While this study offers promising results, some limitations require consideration. Firstly, the research focused on a single case study involving adaptive masking behavior. Additionally, our method of modeling masking behavior focused on providing a simple but effective modification to our ABM that results in complicated dynamics. Our work does not delve into accurately modeling adaptive behaviors. Further, the accuracy of the inferred ODE system depends on the availability and quality of training data. Larger and more diverse datasets are necessary to improve the generalizability of the learned equations. Finally, the inferred ODE system might still be too complex for certain applications. Future work should focus on developing methods for reducing model complexity while maintaining accuracy.

4.2 Future Work

Despite these limitations, several promising avenues for future investigation exist. One promising direction involves extending the framework to different adaptive behaviors relevant to disease spread, such as social distancing or vaccination uptake. Furthermore, more rigorous modeling of the adaptive behavior to capture more accurate and scientifically justified dynamics remain to be done. Since without an accurate model of the adaptive behavior at hand, the recovered dynamics and learned equations will most likely not prove to be useful under real-world scenarios. Our research only provides a simple method of doing this to focus on the equation learning portion of the pipeline thereafter. Additionally, integrating domain knowledge into the equation learning process could guide the discovery of parsimonious and biologically plausible models. Analyzing the learned equations with quantitative methods to gain insights into the underlying mechanisms driving the system's behavior presents another research opportunity. Additionally, we only used one ABM and one compartmental model. Applying this generation to other ABMs with multiple base mathematical models to gain a deeper understanding of the process would likely improve and develop much more insight into this novel process. Finally, creating automated pipelines for generating ABM data, training BINNs, and performing equation learning would make the approach more accessible to a wider range of users.

By addressing the identified limitations and pursuing these promising future directions, we can further enhance the capabilities of equation learning and its application to complex adaptive systems. This will ultimately contribute to the development of improved models for disease prediction and control.

4.3 Conclusions

In conclusion, we successfully obtained an approximate system of ordinary differential equations for an Agent-Based Model with adaptive behaviors. This achievement was made possible by combining neural network approximation with equation learning techniques. The resulting system offers a powerful tool for analyzing and understanding complex adaptive systems, paving the way for further research and applications in various fields.

Author Contributions

Barton, Austin T.: developed code for equation learning, BINN model, training and evaluation, figure generation and analysis, ABM data generation, and repository bookkeeping/organization. Wrote manuscript and generated figures/tables. Klein, Jordan: (FOR JORDAN TO FILL OUT) developed code for modeling masking behavior and figure generation. modeling of masking and figure generation. Flores, Kevin B.: (FOR PROF. KEVIN TO FILL OUT)mentorship, resources, and original idea. Li, Xin: Provided BINNs framework for equation learning on Covasim [?].

Acknowledgements

As part of North Carolina State University's Directed Research for Undergraduates in Mathematics and Statistics (DRUMS) program, our research is supported by the **National Security Agency** (H98230-20-1-0259 and H98230-21-1-0014) and the **National Science Foundation** (DMS#2051010). We would like to thank all of the faculty involved in DRUMS program for the Summer of 2023. We would like to thank Xin Li (Ph.D. student affiliated with Edward P. Fitts, Department of Industrial and Systems Engineering at North Carolina State University) for providing the base mathematical model and programming framework for our project. Lastly, we would like to thank Professor Kevin B. Flores for proposing this research idea, mentorship, and invaluable insights throughout the research process.

References

- [1] Beny Spira. Correlation between mask compliance and covid-19 outcomes in europe. *Cureus*, 2022.
- [2] Erik Rosenstrom, Julie Ivy, Maria Mayorga, Julie Swann, Buse Eylel Oruc, Pinar Keskinocak, and Nathaniel Hupert. High-quality masks reduce covid-19 infections and death in the us. *2021 Winter Simulation Conference (WSC)*, 2021.
- [3] Herbert W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653, 2000.
- [4] Eric Bonabeau. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99(suppl_3):7280–7287, 2002.
- [5] M. R. D'Orsogna, Y. L. Chuang, A. L. Bertozzi, and L. S. Chayes. Self-propelled particles with soft-core interactions: Patterns, stability, and collapse. *Physical Review Letters*, 96(10), 2006.
- [6] Jessica S. Yu and Neda Bagheri. Agent-based models predict emergent behavior of heterogeneous cell populations in dynamic microenvironments. *Frontiers in Bioengineering and Biotechnology*, 8, 2020.
- [7] Melissa Tracy, Magdalena Cerdá, and Katherine M. Keyes. Agent-based modeling in public health: Current applications and future directions. *Annual Review of Public Health*, 39(1):77–94, 2018.
- [8] G. Nigel Gilbert and Klaus G. Troitzsch. *Simulation for the social scientist*. Open Univ. Press, 2011.
- [9] Xinwu Qian, Jiawei Xue, and Satish V. Ukkusuri. Modeling disease spreading with adaptive behavior considering local and global information dissemination, Aug 2020.
- [10] Chad M. Glen, Melissa L. Kemp, and Eberhard O. Voit. Agent-based modeling of morphogenetic systems: Advantages and challenges. *PLOS Computational Biology*, 15(3), 2019.
- [11] Ju-Sung Lee, Tatiana Filatova, Arika Ligmann-Zielinska, Behrooz Hassani-Mahmooei, Forrest Stonedahl, Iris Lorscheid, Alexey Voinov, Gary Polhill, Zhanli Sun, and Dawn C. Parker. The complexities of agent-based modeling output analysis. *Journal of Artificial Societies and Social Simulation*, 18(4), 2015.

- [12] John T. Nardini, Ruth E. Baker, Matthew J. Simpson, and Kevin B. Flores. Learning differential equation models from stochastic agent-based model simulations. *Journal of The Royal Society Interface*, 18(176), 2021.
- [13] Ruth E. Baker and Matthew J. Simpson. Correcting mean-field approximations for birth-death-movement processes. *Physical Review E*, 82(4), 2010.
- [14] Mark A. Chaplain, Tommaso Lorenzi, and Fiona R. Macfarlane. Bridging the gap between individual-based and continuum models of growing cell populations. *Journal of Mathematical Biology*, 80(1–2):343–371, 2019.
- [15] Roberto de Cruz, Pilar Guerrero, Fabian Spill, and Tomás Alarcón. Stochastic multi-scale models of competition within heterogeneous cellular populations: Simulation methods and mean-field analysis. *Journal of Theoretical Biology*, 407:161–183, 2016.
- [16] Nabil T. Fadai, Ruth E. Baker, and Matthew J. Simpson. *Accurate and efficient discretisations for stochastic models providing near agent-based spatial resolution at low computational cost*, 2019.
- [17] John H. Lagergren, John T. Nardini, Ruth E. Baker, Matthew J. Simpson, and Kevin B. Flores. Biologically-informed neural networks guide mechanistic modeling from sparse experimental data. *PLOS Computational Biology*, 16(12), 2020.
- [18] Cliff C. Kerr, Robyn M. Stuart, Dina Mistry, Romesh G. Abeyseuriya, Katherine Rosenfeld, Gregory R. Hart, Rafael C. Núñez, Jamie A. Cohen, Prashanth Selvaraj, Brittany Hagedorn, and et al. Covasim: An agent-based model of covid-19 dynamics and interventions, 2020. *PLOS Computational Biology*, 2020.
- [19] Qiurong Ruan, Kun Yang, Wenxia Wang, Lingyu Jiang, and Jianxin Song. Clinical predictors of mortality due to covid-19 based on an analysis of data of 150 patients from wuhan, china. *Intensive Care Medicine*, 46(5):846–848, 2020.
- [20] D.K. Chu, S. Duda, K. Solo, S. Yaacoub, and H. Schunemann. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of sars-cov-2 and covid-19: A systematic review and meta-analysis. *Journal of Vascular Surgery*, 72(4):1500, 2020.
- [21] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366, 1989.
- [22] Andrew Y. Ng. Feature selection, 1 vs. 1 regularization, and rotational invariance - icml. *Proceedings of the 21st International Conference on Machine Learning*, 2004.

A Supplementary Material

A.1 Code Repository

All code for this project was implemented in Python and is open-source. Code repository can be accessed through this link: https://github.com/abarton51/BINNs_EQL_Covasim.

A.2 Dynamic Tracing Function

We utilized a dynamic tracing function $\chi(t)$ to act as an interacting term with the tracing probabilities in Covasim. Our tracing rate parameter is described as a function of the asymptomatic infected, symptomatic infection, and this χ function, i.e. $\beta(A, Y, \chi)$. There are many different ways to model the tracing rate, but we utilized a piecewise function that was taken from Xin Li's work. The function is described as follows, for each time step of the simulation $t = 1, \dots, t_{max}$, consider

$$\chi(t) = \begin{cases} f_\chi(t/t_{max}) \cdot u_\beta/m & t < 80 \\ u_\beta & \\ f_\chi((t - 40)/t_{max}) \cdot u_\beta/m & t \geq 120 \end{cases} \quad (4)$$

where f_χ is the probability density function for $B(3, 3)$ (B denotes the beta distribution), $t_{max} = 159$, $m \sim B(3, 3)$, and $u_\beta = 0.3$ (tracing rate upper bound).

A.3 Additional Figures and Tables

ABM (Covasim) Hyperparameter	Value/Description
Population Size	500,000
Initial Population Infected	1,000
Duration of Epidemic Simulation	6 Months
Interventions	Testing, Tracing, Masking
Probability of Symptomatic given Infected	0.9
Masking Effectiveness	0.6
Masking Probability Upper Bound	0.75
Masking Probability Lower Bound	0.00
Masking Probability Mean - ω	100
Masking Probability Std - σ	50
Probability of Transmission - Household Layer	0.3
Probability of Transmission - School Layer	0.15
Probability of Transmission - Work Layer	0.25
Probability of Transmission - Community Layer	0.3
Asymptomatic Testing Probability	0.001
Symptomatic Testing Probability	0.1
Quarantined and Symptomatic Testing Probability	0.3
Quarantined and Asymptomatic Testing Probability	0.3
Tracing Probability - Household Layer	1.0
Tracing Probability - School Layer	0.5
Tracing Probability - Work Layer	0.5
Tracing Probability - Community Layer	0.3
Tracing Type	Dynamic Tracing - Piecewise Function

Table 3: Summary of hyperparameters used for our simulated data from Covasim.