

# Deterministic Epidemic Models with Explicit Household Structure

Thomas House, Matt J. Keeling

*Dept. Biological Sciences, University of Warwick, Coventry*

January 17, 2008

## Abstract

For a wide range of airborne infectious diseases, transmission within the family or household is a key mechanism for the spread and persistence of infection. In general, household-based transmission is relatively strong but only involves a limited number of individuals in contact with each infectious person. In contrast, transmission outside the household can be characterised by many contacts but a lower probability of transmission. Here we develop a relatively simple dynamical model that captures these two transmission regimes. We compare the dynamics of such models for a range of household sizes, whilst constraining all models to have equal early growth rate so that all models fit to the same early incidence observations of an epidemic. Finally we consider the use of prophylactic vaccination, responsive vaccination, or antivirals to combat epidemic spread and focus on whether it is optimal to target controls at entire households or to treat individuals independently.

## 1 Introduction

Mathematical modelling is now seen as a key tool in planning for a range of epidemics [1,2] and such models can play an important role in determining an effective control policy [3,4]. Understanding emergent infectious diseases in humans is viewed with increasing importance. Our recent experience with the rapid spread of SARS [5], the perceived threat of bio-terrorism [6] and concerns over influenza pandemics [7] have all highlighted our vulnerability to (re)emerging infections. For all these examples, mathematical modelling has been used to develop an understanding of the relevant epidemiology, as well as to quantify the likely effects of different intervention strategies [2,8–11].

The basis of most epidemiological models is the compartmentalisation of individuals into a range of classes dependent on their infection history. As such the simple *SIR* (Susceptible–Infectious–Recovered) model, which ignores births and deaths, is amongst the most studied [12,13]. Despite its simplicity, this model is generally quite effective at describing the dynamics of a range of infections in many populations [14]. The model is specified by the parameter  $R_0$ , together with initial conditions and the natural time-scale of the infection. These quantities then uniquely determine early growth, peak location and final size of the epidemic.

While such simplicity has its advantages, it precludes capturing the full complexity of population heterogeneities and targeted intervention strategies, both of which have important implications in preparation for potential influenza pandemics or bioterrorism. Faced with these very real threats to human health it is natural to construct parameter-rich simulation models, which aim to capture the full complexity of pathogen transmission and human interaction. While such models have been highly successful [9–11] and are a key requisite for control planning [1, 15, 16], their inherent complexity precludes an intuitive understanding of the wealth of interacting component parameters and variables. A complementary approach, and one adopted here, is to make more modest extensions of the *SIR* model that focus on specific aspects of the complex problem whilst maintaining a relatively small parameter set and deterministic equations.

Here we focus on the transmission dynamics within and between households, essentially constructing a metapopulation model [17], in which contacts that can lead to disease transmission within households are considered to be much stronger than those that can lead to transmission between [18, 19].

There has been much recent interest in household modelling [11, 20, 21]. Formal models dealing with *SIS*-type dynamics in household-structured populations have appeared in [22–24], and for *SIR*- and *SEIR*-type dynamics in [25–27]. Recent work linking more formal results on household models with applied questions, making use of different modelling techniques from ours, has also recently appeared [28, 29].

The present work is unique in presenting analysis of a simple ODE model for transmission of disease that is suited to the analysis of both threshold behaviour and interventions that are affected by transient features of the system. We pay particular attention to robustness of assumptions under reparameterisation of the model, and also consider from the start the need to fit model parameters dynamically to available epidemiological data. In addition to this, we are able to derive analytic expressions for small households that support the robustness of many of our conclusions across the full range of parameter values.

Our model consists of a (relatively) large but simply described set of differential equations, where the stochastic nature of transmission is captured by modelling all possible household configurations. These household models are parameterised to fit a fixed early growth rate, replicating the fitting procedure that would occur from early epidemic data. Finally, we consider control by prophylactic vaccination, responsive vaccination or antiviral treatment [11, 30] and focus on whether such control should be targeted towards entire households or individuals.

## 2 ‘Household gas’ models

Although the *SIR* model is well known and its dynamics have already been studied in considerable detail [12, 13], we define it here both for completeness and to introduce the basic parameters. We consider a closed population, without births or deaths, and adopt a frequentist approach, meaning that  $S$ ,  $I$  and  $R$  are the proportion of the population that are susceptible, infectious and recovered

respectively. The deterministic dynamics of the model are

$$\begin{aligned}\frac{dS}{dt} &= -\beta SI, \\ \frac{dI}{dt} &= \beta SI - \gamma I, \\ \frac{dR}{dt} &= \gamma I\end{aligned}\tag{2.1}$$

where  $\beta$  is the transmission rate and  $\gamma$  is the recovery rate. For this model two key parameters are the basic reproductive ratio,  $R_0$ , and the early growth rate of infection in a naive population.  $R_0$  is defined as the average number of secondary cases produced by a primary infection in a totally susceptible population, and for the simple *SIR* epidemic  $R_0 = \beta/\gamma$ . We consider the early growth of infection to be given by:

$$I(t) = I(0) \exp([r_0 - 1]\gamma t).\tag{2.2}$$

This equation can be used to define the quantity  $r_0$ , which determines the rate of growth during the early stages of an epidemic. For a wide range of unstructured compartmental models (including the simple *SIR* epidemic) these two quantities  $r_0$  and  $R_0$  are identical. Indeed for the model presented in this paper, we find that they are monotonically related, and both equal one at the threshold required to control a disease. Nevertheless, we keep these two quantities separate as they differ numerically and in terms of ease of calculation in the model that follows.

We now consider a population grouped into households. Individuals retain their ‘random’ interaction with the entire population as captured by  $\beta$  in the simple *SIR* epidemic model, but suffer an additional force of infection at rate  $\tau$  per infectious member of their household. To assist in understanding the implications of household structure, we later make the simplifying assumption that all households contain exactly  $n$  individuals—although the methodology readily extends to heterogeneous household sizes. The proportion of households with  $x$  susceptible,  $y$  infectious and  $z$  recovered members is labelled  $H_{x,y,z}$ , with  $\sum_{x,y,z} H_{x,y,z} = 1$ . We stress that  $x$ ,  $y$  and  $z$  are integers referring to numbers of individuals. The proportions of susceptible, infectious and removed individuals in the entire population are then defined as:

$$S := \bar{n}^{-1} \sum_{x,y,z} x H_{x,y,z}, \quad I := \bar{n}^{-1} \sum_{x,y,z} y H_{x,y,z}, \quad R := \bar{n}^{-1} \sum_{x,y,z} z H_{x,y,z},\tag{2.3}$$

where we have defined  $\bar{n} := \sum_{x,y,z} (x + y + z) H_{x,y,z}$ . As mentioned above, for much of this paper we will consider scenarios where for a fixed value of  $n$ , only households obeying  $x + y + z = n$  are present. The relationships in (2.3) allow an informative comparison between the household model and the traditional unstructured models.

The complete dynamics of household types can be determined by considering the rates of transmission between states, reflecting the changes in the number of susceptibles, infected or recovered individuals within households. Rates of change are due to three processes: recovery from infection; within-household transmission; and random transmission between individuals. These terms are grouped together in dynamics of the system as shown below

$$\begin{aligned}\dot{H}_{x,y,z}(t) &= \gamma [-y H_{x,y,z}(t) + (y + 1) H_{x,y+1,z-1}(t)] \\ &\quad + \tau [-xy H_{x,y,z}(t) + (x + 1)(y - 1) H_{x+1,y-1,z}(t)] \\ &\quad + \beta I(t) [-x H_{x,y,z}(t) + (x + 1) H_{x+1,y-1,z}(t)],\end{aligned}\tag{2.4}$$

where  $\tau$  is the within-household transmission rate. Here it should be clear that ‘impossible’ terms are omitted, e.g. for  $z = 0$ , the term in  $H_{x,y+1,z-1}(t)$  should be ignored. Using the summations for the aggregate proportions, the equations for  $SIR$  variables become:

$$\begin{aligned}\frac{dS}{dt} &= -\beta SI - \frac{\tau}{\bar{n}} \sum_{x,y,z} xy H_{x,y,z}(t), \\ \frac{dI}{dt} &= \beta SI - \gamma I + \frac{\tau}{\bar{n}} \sum_{x,y,z} xy H_{x,y,z}(t), \\ \frac{dR}{dt} &= \gamma I,\end{aligned}\tag{2.5}$$

which clearly reduce to (2.1) when either  $\tau = 0$  or if the household size  $n = 1$ . We note that, as for systems such as (2.1), the equations (2.5) represent the large population limit of a system where mixing outside the household, represented by the rate  $\beta$ , is homogeneous.

## 2.1 Early dynamics

To study the early dynamics of our household model, we need to define two important distributions of household types.  $\widehat{H}_{x,y,z}$  is given by the eigenvector associated with the dominant eigenvalue,  $\Lambda_1 = (r_0 - 1)\gamma$ , of the system 2.4 linearised about the disease-free equilibrium. We consider the early period of an epidemic, during which the proportion of susceptibles is still close to 1. In a population of households of size  $n (= x + y + z)$ , the household proportions during this phase of the epidemic are given by

$$H_{x,y,z}(t) = \delta_{0,y}\delta_{0,z} + \widehat{H}_{x,y,z}I(t), \quad \text{where } I(t) \sim e^{\Lambda_1 t}.\tag{2.6}$$

Throughout  $\delta_{x,y}$  is the Kronecker delta, which is one if  $x$  and  $y$  are identical and zero otherwise. Realistically,  $\widehat{H}_{x,y,z}$  can only be determined numerically for  $n > 2$ , however when  $n = 2$ , so that all individuals are in households of size two, it is possible to calculate these terms analytically, giving

$$r_0(n = 2) = \frac{\beta}{\gamma} + \frac{\beta + \tau}{2\gamma} \left[ \left( 1 + \frac{4\beta\tau}{(\beta + \tau)^2} \right)^{1/2} - 1 \right].\tag{2.7}$$

While in practise this expression can be derived simply from consideration of the dynamical equations for  $H_{1,1,0}(t)$  and  $I(t)$  during the early states of an epidemic, for completeness we provide the full distribution of households for this system in appendix A. Although  $r_0$  cannot be found in general for  $n > 2$ , from first principles we find that when within household transmission is rapid ( $\tau \rightarrow \infty$ ) we have  $r_0 \rightarrow n\beta/\gamma$ , since a single introduction into a household leads instantly to  $n$  newly infectious individuals. This value therefore acts as an upper bound for the general case.

An additional distribution of households that is highly informative is the distribution of ‘final epidemic’ sizes within a singly infected, isolated household, which we call  $H_{x,z}^*$ . Formally, this can be defined as

$$\begin{aligned}H_{x,z}^* &:= \lim_{\substack{\beta \rightarrow 0 \\ t \rightarrow \infty}} \left[ H_{x,0,z}(t) \right], \\ \text{where } H_{x,y,z}(0) &= \delta_{x,n-1}\delta_{y,1}\delta_{z,0}.\end{aligned}\tag{2.8}$$

By solving an appropriate Markov process, or via the approach of [31, 32], this distribution can be found analytically, so that for households of size 2

$$H_{2,0}^* = 0, \quad H_{0,2}^* = \frac{\tau}{\gamma + \tau}, \quad H_{1,1}^* = \frac{\gamma}{\gamma + \tau}. \quad (2.9)$$

### 3 Parameter fitting

For the simple *SIR* epidemic of equation (2.1), assuming that the initial proportion of the population infected is sufficiently small to ignore, we have only two parameters: the recovery rate  $\gamma$  sets the basic time-scale of the dynamics, while  $R_0 = \beta/\gamma$  determines all other aspects, including the early epidemic growth rate, the final size of the epidemic and the levels of control required for eradication. It is generally assumed that  $\gamma$ , or more precisely the average infectious period  $1/\gamma$ , can be estimated from detailed observations of infectious individuals. Therefore, for the simple *SIR* epidemic, the early growth rate of infected cases allows complete parameterisation of the model. This is not the case for the household model where we have two free parameters,  $\beta$  and  $\tau$ . We therefore consider three assumptions that lead to just one free parameter.

Throughout, we consider parameters consistent with influenza-like illness (as outlined in [2] and references therein). For all three assumptions, we consider  $r_0$  to be observed, and assume in the numerical results that  $r_0 = 2$ . We then show how this impacts on the parameters  $\beta$  and  $\tau$ .

#### 3.1 Assumption 1: Fixed within-household transmission rate

The simplest assumption is that the within-household transmission rate  $\tau$  has been observed directly. Most often this is achieved by monitoring the number of secondary infections within a household that can be assumed to be relatively isolated from other sources of infection [33]. We fixed  $\tau$  by making the assumption that there is an 80% probability of infection being passed between two members of the same household, which leads to  $\tau = 4\gamma$ .

The population-level transmission rate,  $\beta$ , is then parameterised to achieve the desired growth parameter,  $r_0$ . For large household sizes ( $n \geq 3$ ) this parameterisation must be achieved by calculating  $r_0$  numerically from the eigenvalue approach, although for  $n = 2$  analytical results are again possible:

$$\beta(n = 2) = r_0 \gamma \frac{\gamma r_0 + \tau}{(\gamma r_0 + 2\tau)}. \quad (3.1)$$

So, for the fixed parameters used ( $r_0 = 2$  and  $\tau = 4\gamma$ ) we find that  $\beta(n = 2) = 6\gamma/5$ .

#### 3.2 Assumption 2: Fixed ratio of within-household to between-household transmission rates

A second approach is to fit both  $\tau$  and  $\beta$  so that their ratio remains constant. This assumption reflects the concept that we know how much greater transmission is within the household compared to between households. To maintain consistency with the above assumption when  $n = 2$  we set  $\beta/\tau = 0.3$ , giving a household transmission rate more than 3 times the general transmission rate in the population.

### 3.3 Assumption 3: Fixed proportion of early infections happening within households

The proportion of infections that happen within the household varies throughout the course of an epidemic, however it is a quantity that can be measured and would be ascertainable from the detailed contact tracing that is performed in the early stages of an epidemic. If  $P$  is the proportion of infections that occur due to household-based transmission during the early stages of the epidemic then

$$\beta = (1 - P)\gamma r_0, \quad \tau = \frac{\bar{n}\gamma r_0 P}{\sum_{x,y,z} xy \hat{H}_{x,y,z}}. \quad (3.2)$$

This expression is derived in appendix B from the early behaviour of the system. For consistency we set  $P = 40\%$ , which leads to the same  $\beta$  and  $\tau$  values when  $n = 2$  as in the other assumptions.

## 4 Epidemic dynamics

We now use each fitting method in turn, together with numerical integrations of the household model (2.4), to investigate the implications of household structure for both the shape of the infection curves and summary statistics of the epidemic.

Figure 1 (top row) shows the changing prevalence of infection through the epidemic for ‘pure’ households of size  $n$ . All epidemics are constrained to have the same early growth by matching  $r_0$ , and we additionally fit the initial, very small, level of infection such that the early epidemic curves overlap. Given these tight constraints on the early growth, not surprisingly the ensuing epidemics are relatively similar. The second row of figure 1 highlights the differences by subtracting the simple *SIR* epidemic curve ( $n = 1$ ) from the household-based results. We consistently find that adding household structure leads to a more sustained epidemic phase (density-dependent saturation is weaker) and a more rapid decline after the epidemic peak. We note that it is only for fitting assumption 2 that the changes are monotonic with respect to  $n$ ; for the other assumptions the difference between the household and simple *SIR* epidemic models is maximised at intermediate  $n$ . The bottom two rows of figure 1 exemplify these changes, plotting the peak epidemic size and the total proportion of the population ever infectious (final epidemic size) for household sizes from 1 to 10. Clearly parameterisation assumption 2 admits the greatest epidemic changes as the household size varies. Finally, the central row of this figure gives the values of  $\tau$  and  $\beta$  that are used in the simulations and highlight the impact of the three assumptions.

## 5 Modelling control measures

In this section we consider modelling three possible control methods: prophylactic vaccination, responsive vaccination and distribution of antivirals. These methods operate in very different ways: prophylactic vaccination suppresses infection by reducing the initial number of available susceptibles in the population; responsive vaccination removes susceptibles at a certain rate during an epidemic; and antivirals are administered to infected individuals in order to reduce their subse-

quent transmission. For all of these control options a key issue is whether individuals or entire households should be the focus of intervention.

## 5.1 Prophylactic Vaccination

Prophylactic vaccination (before an epidemic begins) essentially changes the initial distribution of susceptible and immune individuals within households which is taken as the initial conditions for the epidemic. For a vaccination level of  $p$  we examine two scenarios: firstly household vaccination, with initial conditions

$$H_{n,0,0}(0) = 1 - p, \quad H_{0,0,n}(0) = p, \quad (5.1)$$

for households of size  $n$  and secondly individual-level vaccination, with initial conditions

$$H_{n-z,0,z}(0) = \frac{n!}{(n-z)!z!} p^z (1-p)^{n-z}. \quad (5.2)$$

This latter expression represents a fully random distribution of vaccine through the population, regardless of household structure. We note that both household and individual vaccination as described above are extreme cases, and that for realistic strategies the situation will be further complicated by additional units of population structure such as workplaces and social groups. Nevertheless, consideration of these two strategies will provide a useful intuitive insight into how to target interventions on the basis of population structure.

A considerable body of work exists on vaccination thresholds (recently [34] and also references therein) that deals with this issue at a high level of generality. We present here, for completeness, the standard method for obtaining thresholds for control in household-structured populations presented in these references. This more formal approach will give results compatible with our numerical results, which we obtain by finding intervention parameter thresholds at which  $r_0$  falls below 1.

For the two scenarios under consideration is it simple to search numerically for the value of  $p$  that generates a zero growth rate. However, a more analytically tractable approach is to consider the expected number of cases in a singly infected household (c.f.  $H_{x,z}^*$  defined in (2.8)). We define  $Z_\infty(x)$  to be the expected total number of cases in a household which initially has  $x$  susceptibles and one infectious individual—note that the household size  $n$  does not explicitly feature in this quantity, so

$$Z_\infty(x) = \sum_z z H_{x-z,z}^*. \quad (5.3)$$

For  $x = 1, 2, 3$ , an explicit formula for  $Z_\infty$  is given by

$$Z_\infty(1) = 1, \quad Z_\infty(2) = \frac{2\tau + \gamma}{\tau + \gamma}, \quad Z_\infty(3) = \frac{6\tau^3 + 13\tau^2\gamma + 6\tau\gamma^2 + \gamma^3}{(\tau + \gamma)^2(2\tau + \gamma)}. \quad (5.4)$$

We can now determine the household-level basic reproductive ratio,  $R_*$ , (defined as the average number of secondary households infected by a single infected household in a totally susceptible population) using the method of [27]. When considering uncontrolled epidemics or prophylactic vaccination  $R_*$  is given by

$$R_* = \frac{\beta}{\gamma} \sum_{x,y,z} x H_{x,y,z}(0) Z_\infty(x) \quad (5.5)$$

This reproductive ratio clearly depends on the initial distribution of household types,  $H_{x,y,z}(0)$ , as well as the expected number of external secondary cases generated,  $\beta/\gamma$ . The critical eradication threshold then occurs when  $R_* = 1$ .

For pairs (i.e. households of size two) this method can be checked against a still simpler, analytic approach from first principles, leading to a critical vaccination level of

$$p_{\text{crit}}(n = 2) = \begin{cases} 1 - \frac{\tau + \gamma}{2\tau} \left[ \left( 1 + \frac{2\gamma\tau}{\beta(\tau + \gamma)} \right)^{1/2} - 1 \right] & \text{for individual-based vaccination,} \\ 1 - \frac{\gamma(\tau + \gamma)}{2\beta(2\tau + \gamma)} & \text{for household-based vaccination.} \end{cases} \quad (5.6)$$

These analytic results are the same as those obtained numerically, for zero early growth rate, as shown in figure 2 for the three different fitting assumptions. These show a consistent threshold benefit to targeting individuals rather than households when vaccinating. This result is not surprising if one considers the number of strong within-household transmission links broken by a given vaccination strategy. When choosing to vaccinate a single individual, it is clearly optimal to immunise an individual in a large susceptible household as this individual is both most at risk of infection and also most able to transmit infection. Once such an individual has been vaccinated, this weakens the case for immunising the rest of the household, which then represents a smaller pool of susceptibles. We note that this result, readily obtained from our model, is consistent with the more formal results of [27, 35].

## 5.2 Responsive vaccination

For some diseases, vaccines are long-lasting and distributed well in advance of any outbreak, while for others vaccination would need to take place at the start of or during the epidemic. This leads to a situation where susceptibles are removed at a rate  $V(S)$  so that the simple *SIR* epidemic (2.1) becomes

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI - V(S), \\ \frac{dI}{dt} &= \beta SI - \gamma I, \\ \frac{dR}{dt} &= \gamma I + V(S). \end{aligned} \quad (5.7)$$

While in practice  $V$  will be a function of  $S$  reflecting the increasing difficulty of successfully discovering and vaccinating susceptibles as the number of them reduces, we make the simplifying assumption that  $V(S) = v\Theta(S)$ , where  $v$  is a constant and  $\Theta$  is the step function, leaving the rate of vaccination constant for all  $S > 0$ .

Only an infinite value for  $v$  makes complete control of the disease possible. Our focus of investigation is therefore on the final size of the epidemic as a function of  $v$ , the results for which are shown in figure 2. For any household size, for sufficiently low values of  $v$  the final size will be close to that without vaccination, while for sufficiently high values of  $v$  the final size will approach the initial number of cases,  $I_0$ . On the first row of figure 2 we show the results for the system (5.7), which we label  $n = 1$ , for initial conditions  $I_0 = 10^{-3}, 10^{-6}$ . The subsequent rows show the difference between different household sizes and the  $n = 1$  system, given different targets for the vaccination.



We consider three such targets. The first of these is *individual*, where we randomly vaccinate individuals for vaccination at a rate  $v$  in a way that recovers the form of (5.2) in the infinite  $v$  limit. We also consider two household-based targets, which both recover the form of (5.1) in the infinite  $v$  limit. These two methods differ, however, in the benefit attributed to rate of vaccination from household targeting, and are called the *slow household* and *fast household* methods respectively

This gives us dynamics as below:

$$\begin{aligned} \dot{H}_{x,y,z}(t) = & \gamma [-yH_{x,y,z}(t) + (y+1)H_{x,y+1,z-1}(t)] \\ & + \tau [-xyH_{x,y,z}(t) + (x+1)(y-1)H_{x+1,y-1,z}(t)] \\ & + \beta I(t) [-xH_{x,y,z}(t) + (x+1)H_{x+1,y-1,z}(t)] \\ & + \frac{v}{S(t)} \times \begin{cases} -xH_{x,y,z}(t) + (x+1)H_{x+1,y,z-1}(t) & \text{for individual,} \\ -(1-\delta_{x,0})H_{x,y,z}(t) + \delta_{x,0} \sum_{\xi} H_{\xi,y,z-\xi}(t) & \text{for slow household,} \\ (x+y+z) \left[ -(1-\delta_{x,0})H_{x,y,z}(t) + \delta_{x,0} \sum_{\xi} H_{\xi,y,z-\xi}(t) \right] & \text{for fast household.} \end{cases} \end{aligned} \quad (5.8)$$

We note that both individual and slow household vaccination preserve the relation  $\dot{S}(t) = \dots - v$ , however for the fast household we assume that there is maximal benefit in terms of rate from targeting households, i.e. that all household sizes can be vaccinated at the same rate. While this is clearly an extreme case, there will be some scenarios in which there is a rate benefit from targeting households that will lie somewhere between the fast and slow household cases.

The results for these intervention methods are displayed in the bottom two rows of figure 3, where it can be seen that the main benefits from adopting any strategy lie in the intermediate values of  $v$ . The fast household is always preferable to individual targeting, which is in turn always preferable to the slow household. This means that in terms of mitigation of the overall effects of an epidemic (measured by the total number infected) if there is any benefit in terms of the rate at which vaccine can be distributed when households are targeted then calculations of optimum vaccination strategy for pre-epidemic vaccination thresholds cannot simply be extrapolated to the dynamic case.

### 5.3 Anti-viral treatment

Anti-viral drugs can be used both as short-term prophylaxis (to prevent infection) and as treatment [11,30]. To model the former requires the addition of extra compartments, since protection offered is temporary and does not lead to immunity. Such additional complexity is outside the scope of the current model; we ignore the effect of anti-viral drugs on susceptible individuals, and concentrate on their role in the treatment of disease. We assume that the only effect from antiviral drugs is to cause infectious individuals to recover. We therefore make an optimistic assumption about the speed with which antivirals work, but a pessimistic assumption regarding their impact on susceptibles. At the household level these assumptions are largely irrelevant since the removal of infected individuals halts within-household transmission.

We call the rate at which infectious individuals are detected the *antiviral rate*. After detection, anti-viral drugs are given either to the individual, in which case we write  $\alpha$  for the antiviral rate, or to the entire household, when we write  $a$  for the antiviral rate. When antivirals are given to

individuals, then  $\alpha$  basically acts to increase the effective recovery rate ( $\gamma$  is replaced by  $\gamma + \alpha$ ); however for treatment of entire households the dynamics are more complex, leading to extra terms in the dynamics, so that

$$\begin{aligned} \dot{H}_{x,y,z}(t) &= (\gamma + \alpha) [-yH_{x,y,z}(t) + (y + 1)H_{x,y+1,z-1}(t)] \\ &\quad + \tau [-xyH_{x,y,z}(t) + (x + 1)(y - 1)H_{x+1,y-1,z}(t)] \\ &\quad + \beta I(t) [-xH_{x,y,z}(t) + (x + 1)H_{x+1,y-1,z}(t)] \\ &\quad - ayH_{x,y,z}(t) + \delta_{0,y} \sum_{y'} ay' H_{x,y',z-y'}(t), \end{aligned} \quad (5.9)$$

We define the critical value of an antiviral rate to be the value at which it leads to an  $r_0$  value of 1. This critical value will clearly be lower for household-based intervention, as detection time is effectively decreased for individuals in households with multiple infections. We note that in all our results, we set either  $\alpha$  or  $a$  to be zero and consider the effect of the other in isolation for simplicity.

Again it is possible to produce exact results for households of size  $n = 2$ . For individual based treatment where households consist of pairs, the critical detection rate is:

$$\alpha_{\text{crit}}(n = 2) = \gamma(r_0 - 1). \quad (5.10)$$

Analytic results for household targeting at  $n = 2$  and individual targeting at  $n = 3$  are given in Appendix C. Although these are too complex to provide an intuitive understanding, they do confirm the numerical results, and demonstrate that they are qualitatively robust under reparameterisation of the model.

Clearly, household targeting of antiviral treatment decreases the rate at which cases must be detected, however at the same time it introduces a wastage of drugs on individuals who are not infected at the time the drugs are administered. While the short-term prophylactic effect that this treatment confers will be of some benefit, it is still reasonable to assume that there will be a practical trade-off between the desire to administer drug treatment quickly and the desire to treat the sick. We define two quantities that capture the relevant features of household-based intervention. The first of these we call the *household-based rate benefit*,  $\phi$ , which quantifies the benefits of household targeting in terms of the relative decrease of the critical rates necessary to control the epidemic.

$$\phi(n) := \frac{\alpha_{\text{crit}}(n) - a_{\text{crit}}(n)}{\alpha_{\text{crit}}(n)}, \quad (5.11)$$

where as above we write  $\alpha$  for the individual-targeted antiviral rate and  $a$  for the household-targeted antiviral rate. The second quantity we consider is the rate at which antiviral drugs are distributed to non-infectious individuals,  $w$ . This will be given in general by

$$w(t) := a \sum_{x,y,z} (x + z)yH_{x,y,z}(t). \quad (5.12)$$

During the early phase of a disease outbreak that is successfully controlled by household-targeted antiviral treatment at the critical rate, a normalised constant can be defined, which we call the *Household-based wastage rate*, given by

$$W := \frac{w(t)}{a_{\text{crit}}\bar{n}^2 I(t)} = \frac{1}{\bar{n}^2} \sum_{x,y,z} (x + z)y\hat{H}_{x,y,z}. \quad (5.13)$$

Both of these quantities are normalised to obey  $\phi = W = 0$  for populations without household structure and to lie in the range  $[0, 1)$  in general. For our parameter choices, the relationship between these quantities is shown in figure 4. These graphs show that it is typically not possible to increase the rate benefit without adversely affecting the wastage rate. In addition, the effects of household-based intervention are significantly worse in both senses for our third fitting assumption, which happens due to the relatively low level of within-household transmission under that assumption as the household size grows. This leads to a situation where infectious individuals are more evenly distributed throughout households making it more likely that antivirals will be wasted if a household is targeted.

## 6 Discussion

In this paper, we have developed a relatively simple model for the transmission of an infectious agent through a population of well mixed households. To make a fair comparison as household size is varied, we fix the early epidemic growth rate—essentially all models would agree with the case-report data gathered during the first few generations of an epidemic. We feel that this agreement is vital if we are to compare models reliably. However, even with this constraint there is still a free parameter and three different assumptions are utilised depending upon the type of information that is likely to be available.

Our results show that the general epidemic profile is fairly consistent as household size varies, although parameterisation assumption 2 (where the ratio of within-household to between-household transmission is fixed) shows the greatest variation.

The results for control are more significant. In the absence of other risk factors, vaccinating individuals randomly is better than targeting entire households, with the critical level of vaccination needed for eradication depending on both the household size and the fitting assumption. In fact these results are consistent with earlier results, such as [27], that leaving a homogeneous number of susceptibles in each household following vaccination is optimal, although this may be impractical. When considering antiviral treatment interpretation of the results is more complex: although treating entire households requires a lower rate of detection to eradicate the epidemic, some antiviral treatment is wasted on uninfected individuals.

The models used in this paper are clearly an abstraction of the real dynamics of infectious diseases. We have only dealt with results from populations where households and individuals are identical; in reality households vary in size and in composition and individuals vary in their interactions both within and between households. In addition, we have assumed a simple Markovian *SIR* model for the infection dynamics. In reality, most diseases obey the *SEIR* paradigm, which adds an exposed class through which infected individuals pass before becoming infectious. It is also significantly more realistic to have non-exponential waiting times in the exposed and infectious classes—different assumptions about the distribution of waiting times is known to have a significant impact on the success of control [36]. To incorporate such additional heterogeneities requires a considerable degree more model complexity. Instead we have highlighted some general concepts for epidemic modelling in populations with household structure that can provide an intuitive understanding, although we note that more complex simulation models should be considered when

determining policy.

## Acknowledgements

This work was funded by EU grant INFTRANS (FP6 STREP; contract no. 513715). We give thanks to members of the Infectious Disease Epidemiology Research Interest Group at Warwick, to Joshua Ross, and to other members of the INFTRANS consortium for their helpful comments on this work.

## References

- [1] N.M. Ferguson, M.J. Keeling, Edmunds W.J., R. Gani, B.T. Grenfell, R.M. Anderson, and S.A. Leach. Planning for smallpox outbreaks. *Nature*, 425:681–685, 2003.
- [2] N.M. Ferguson, D.A.T. Cummings, C. Fraser, J.C. Cajka, P.C. Cooley, and D.S. Burke. Strategies for mitigating an influenza pandemic. *Nature*, 442(7101):448–452, April 2006.
- [3] M.J. Keeling, M.E.J. Woolhouse, D.J. Shaw, L. Matthews, M. Chase-Topping, D.T. Haydon, S.J. Cornell, J. Kappey, J. Wilesmith, and B.T. Grenfell. Dynamics of the 2001 UK foot and mouth epidemic: Stochastic dispersal in a heterogeneous landscape. *Science*, 294:813–817, 2001.
- [4] N.M. Ferguson, C.A. Donnelly, and R.M. Anderson. The foot-and-mouth epidemic in Great Britain: Pattern of spread and impact of interventions. *Science*, 292:1155–1160, 2001.
- [5] C.A. Donnelly, A.C. Ghani, G.M. Leung, A.J. Hedley, C. Fraser, S. Riley, L.J. Abu-Raddad, L.M. Ho, T.Q. Thach, P. Chau, K.P. Chan, T.H. Lam, L.Y. Tse, T. Tsang, S.H. Liu, J.H.B. Kong, E.M.C. Lau, N.M. Ferguson, and R.M. Anderson. Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. *Lancet*, 361:1761–1766, 2003.
- [6] H.C. Lane, J.L. Montagne, and A.S. Fauci. Bioterrorism: a clear and present danger. *Nature Med.*, 7:1271–1273, 2001.
- [7] World Health Organisation. Epidemic and pandemic alert and response (EPR)—avian influenza. [http://www.who.int/csr/disease/avian\\_influenza/](http://www.who.int/csr/disease/avian_influenza/).
- [8] S. Riley, C. Fraser, C.A. Donnelly, A.C. Ghani, L.J. Abu-Raddad, A.J. Hedley, G.M. Leung, L.M. Ho, T.H. Lam, T.Q. Thach, P. Chau, K.P. Chan, P.Y. Leung, T. Tsang, W. Ho, K.H. Lee, E.M.C. Lau, N.M. Ferguson, and R.M. Anderson. Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of public health interventions. *Science*, 300:1961–1966, 2003.
- [9] M.E. Halloran, I.M. Longini, A. Nizam, and Y. Yang. Containing bioterrorist smallpox. *Science*, 298(1428-1432), 2002.

- [10] S. Riley and N.M. Ferguson. Smallpox transmission and control: Spatial dynamics in Great Britain. *Proc. Nat. Acad. Sci. USA*, 103:12637–12642, 2006.
- [11] N.M. Ferguson, D.A.T. Cummings, S. Cauchemez, C. Fraser, S. Riley, A. Meeyai, S. Iam-sirithaworn, and D.S. Burke. Strategies for containing an emerging influenza pandemic in southeast Asia. *Nature*, 437:209–214, 2005.
- [12] W.O. Kermack and A.G. McKendrick. Contribution to the mathematical theory of epidemics. *Proc. Roy. Soc. Lond. A*, 115:700–721, 1927.
- [13] R.M. May and R.M. Anderson. Population biology of infectious diseases part II. *Nature*, 280:455–461, 1979.
- [14] R.M. Anderson and R.M. May. *Infectious diseases of humans*. Oxford University Press, 1992.
- [15] M.J. Keeling. Models of foot-and-mouth disease. *Proc. Roy. Soc. Lond. B*, 272:1195–1202, 2005.
- [16] S. Riley. Large-scale spatial-transmission models of infectious disease. *Science*, 316:1298–1301, 2007.
- [17] M.J. Keeling, O.N. Bjørnstad, and B.T. Grenfell. Metapopulation dynamics of infectious diseases. In I. Hanski and O.E. Gaggiotti, editors, *Ecology Genetics and Evolution of Metapopulations*, pages 415–446. Elsevier Academic Press, 2004.
- [18] I.M. Longini, J.S. Koopman, A.S. Monto, and J.P. Fox. Estimating household and community transmission parameters for influenza. *Am. J. Epi.*, 115:736–751, 1982.
- [19] I.M. Longini and J.S. Koopman. Household and community transmission parameters from final distributions of infections in households. *Biometrics*, 38:115–126, 1982.
- [20] N.G. Becker and K. Dietz. The effect of household distribution on transmission and control of highly infectious-diseases. *Math. BioSci.*, 127:207–219, 1995.
- [21] J.T. Wu, S. Riley, C. Fraser, and G.M. Leung. Reducing the impact of the next influenza pandemic using household-based public health interventions. *PLoS Medicine*, 3:1532–1540, 2006.
- [22] Frank Ball. Stochastic and deterministic models for SIS epidemics among a population partitioned into households. *Math. BioSci.*, 156:41–67, 1999.
- [23] G. Ghoshal, L.M. Sander, and I.M. Sokolov. SIS epidemics with household structure: the self-consistent field method. *Math. BioSci.*, 190:71–85, 2004.
- [24] Peter Neal. Stochastic and deterministic analysis of SIS household epidemics. *Adv. Appl. Prob.*, 38:943–968, 2006.
- [25] Frank Ball, Philip D. O’Neill, and James Pike. Stochastic epidemic models in structured populations featuring dynamic vaccination and isolation. *J. Appl. Prob.*, 4:571–585, 2007.

- [26] Frank Ball and Owen Lyne. Epidemics among a population of households. In *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods and Theory*, volume 126 of *IMA Volumes in Mathematics and its Applications.*, pages 115–142. Springer–Verlag, 2002.
- [27] F. Ball, D. Mollison, and G. Scalia-Tomba. Epidemics with two levels of mixing. *Ann. Appl. Probab.*, 7(1):46–89, 1997.
- [28] P.J. Dodd and N.M. Ferguson. Approximate disease dynamics in household-structured populations. *J. R. Soc. Interface*, 4:1103–1106, 2007.
- [29] Christophe Fraser. Estimating individual and household reproduction numbers in an emerging epidemic. *PLoS ONE*, 2(8):1–12, 2007.
- [30] M.E. Halloran, F.G. Hayden, Y. Yang, I.M. Longini, and A.S. Monto. Antiviral effects on influenza viral transmission and pathogenicity: Observations from household-based trials. *Am. J. Epi.*, 165:212–221, 2007.
- [31] Frank Ball. A unified approach to the distribution of total size and total area under the trajectory of infectives in epidemic models. *Adv. Appl. Prob.*, 18(2):289–310, 1986.
- [32] Norman T. Bailey. *The Mathematical Theory of Infectious Diseases and its Applications*. Griffin, London, 1975.
- [33] R.E. Hope-Simpson. Infectiousness of communicable diseases in the household. *The Lancet*, 2:549–554, 1952.
- [34] Frank Ball and Owen Lyne. Optimal vaccination schemes for epidemics among a population of households, with application to variola minor in Brazil. *Stat. Methods Med. Res.*, 15(5):481–497, 2006.
- [35] F. Ball and O. Lyne. Optimal vaccination policies for stochastic epidemics among a population of households. *Mathematical Biosciences*, 177–178:222–354, 2002.
- [36] H.J. Wearing, P. Rohani, and M.J. Keeling. Appropriate models for the management of infectious diseases. *PLoS Medicine*, 2:621–627, 2005.

## A Dominant eigenvector for $n = 2$

For a population of pairs, the system's dominant eigenvalue is given by

$$\begin{aligned}
\widehat{H}_{2,0,0} &= \frac{4\beta}{-B - 2\beta + 2\gamma + \tau}, \\
\widehat{H}_{1,1,0} &= \frac{B - 2\beta - \tau}{2\tau}, \\
\widehat{H}_{1,0,1} &= \frac{\gamma(-4\beta^2 + 2(B + \gamma - 3\tau)\beta + \gamma(\tau - B))}{2\tau(\gamma(\gamma + \tau) - 2\beta(\gamma + 2\tau))}, \\
\widehat{H}_{0,1,1} &= -\frac{\gamma(\tau^2 + (B - 8\beta - 3\gamma)\tau + (B - 2\beta)(2\beta + \gamma))}{2\tau(\gamma(2\beta + \gamma) - (4\beta + \gamma)\tau)}, \\
\widehat{H}_{0,2,0} &= \frac{(B - 2\beta)(2\beta + \gamma) - (6\beta + \gamma)\tau}{2\gamma(2\beta + \gamma) - 2(4\beta + \gamma)\tau}, \\
\widehat{H}_{0,0,2} &= -\frac{4\gamma^2\tau}{-8\beta^3 - 4(B + 5\tau)\beta^2 + 2(\gamma^2 - \tau(2B + \tau))\beta + (B - \tau)\tau^2 + \gamma^2(B + 3\tau)},
\end{aligned} \tag{A.1}$$

where

$$B := \pm\sqrt{4\beta^2 + 12\tau\beta + \tau^2}. \tag{A.2}$$

## B Proportion of infections within the household

Inserting (2.6) into (2.5) at linear order in  $I$  gives

$$\begin{aligned}
\dot{I}(t) &= (\beta + K - \gamma)I(t), \quad \text{for} \\
K &:= \frac{\tau}{n} \sum_{x,y,z} xy\widehat{H}_{x,y,z}.
\end{aligned} \tag{B.1}$$

Clearly, then, during the early growth phase of an epidemic, the proportion of infections happening within-household will be

$$P = \frac{K}{K + \beta}. \tag{B.2}$$

At the same time,

$$r_0 = \frac{\beta}{\gamma} + \frac{K}{\gamma}. \tag{B.3}$$

Stipulating that  $r_0$  and  $P$  should remain constant for different household distributions gives condition (3.2).

## C Analytical antiviral results for small households

### C.1 Critical individual-targeted antiviral rate for $n = 3$

For triples, it is not possible to solve for the critical individual antiviral rate exactly, however we may take the Ansatz

$$\alpha_{\text{crit}}(n = 3) = \gamma(r_0 + \epsilon - 1), \tag{C.1}$$

and solve the equation  $R_*(n=3) = 1$  to linear order in  $\epsilon$ . This gives the following solution

$$\begin{aligned}\epsilon &= \frac{\frac{C\gamma r_0}{A\beta(n=3)} - 1}{\frac{\gamma B}{A} - \gamma D - \frac{1}{r_0}}, \quad \text{for} \\ A &= 6\tau^3 + 13\tau^2\gamma r_0 + 6\tau\gamma^2 r_0^2 + \gamma^3 r_0^3, \\ B &= 13\tau^2 + 12\tau\gamma r_0 + 3\gamma^2 r_0^2, \\ C &= (\tau + \gamma r_0)^2(2\tau + \gamma r_0), \\ D &= \frac{5\tau + 3\gamma r_0}{(\tau + \gamma r_0)(2\tau + \gamma r_0)}.\end{aligned}\tag{C.2}$$

## C.2 Critical household-targeted antiviral rate for $n = 2$

For the case of household-targeted antiviral treatment, the equation (5.5) no longer holds, and instead we must calculate

$$R_* = \beta \sum_{x,y,z} H_{x,y,z}(0) \int_0^\infty Y(x;t) dt, \tag{C.3}$$

where  $Y(x;t)$  is the expected number of infectious individuals at time  $t$  in a clique with one initial infection and  $x - 1$  initial susceptibles. Clearly, for a process where  $\dot{R}(t) = \gamma I(t)$ , the expressions (C.3) and (5.5) will be equivalent, however for household-targeted antiviral treatment we have I-I pairs ‘recovering’ directly to R-R pairs at a rate  $2a$  rather than via I-R pairs at rate  $2\gamma$ . Performing the full calculation gives

$$\begin{aligned}a_{\text{crit}}(n=2) &= \frac{1}{3} \left( \beta - 3\gamma - \tau + A^{1/3} + (\beta^2 + \tau\beta + \tau^2)A^{-1/3} \right), \quad \text{where} \\ A &= \beta^3 + \frac{3}{2}\tau\beta^2 - \frac{3}{2}\tau(\tau - 9\gamma)\beta - \tau^3 \\ &\quad + 3\sqrt{\frac{3}{2}\beta\tau \left[ \frac{1}{2}(4\gamma - \tau)\beta^3 + (3\gamma - \tau)\tau\beta^2 + \frac{1}{2}(27\gamma^2 - 6\tau\gamma - \tau^2)\tau\beta - 2\gamma\tau^3 \right]}.\end{aligned}\tag{C.4}$$



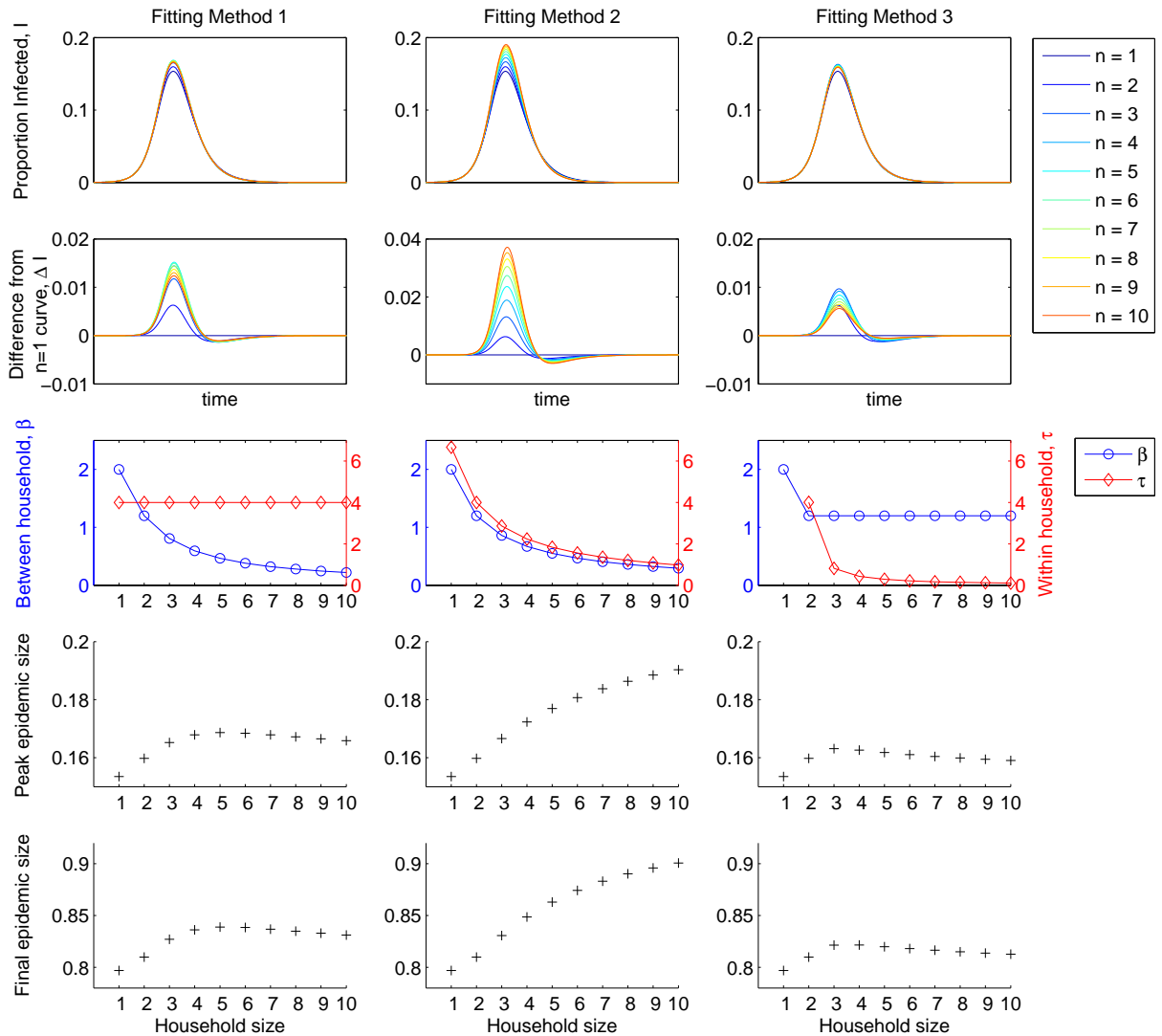


Figure 1: Numerical results for the household model, for three different fitting methods. The first row shows infection curves with the same early growth for different values of the ‘pure’ household size  $n$ , and on the row beneath are the differences between those curves and the simple *SIR* epidemic ( $n = 1$ ) curve. The middle row shows how the between-household transmission rate  $\beta$  and the within-household rate  $\tau$  are scaled over different household sizes. The final two rows show how the peak number of infectious individuals and the final epidemic size are influenced by household size.

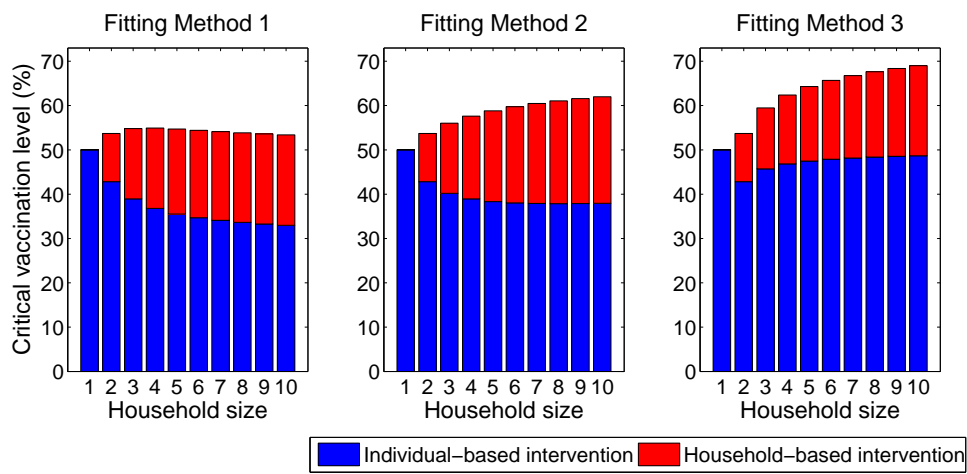


Figure 2: Effects of prophylactic vaccination. The critical percentage of the population that needs to be immune prior to the epidemic to prevent an outbreak is shown for three fitting methods and for a range of ‘pure’ household sizes and for vaccination targeted either at the household or individual level.

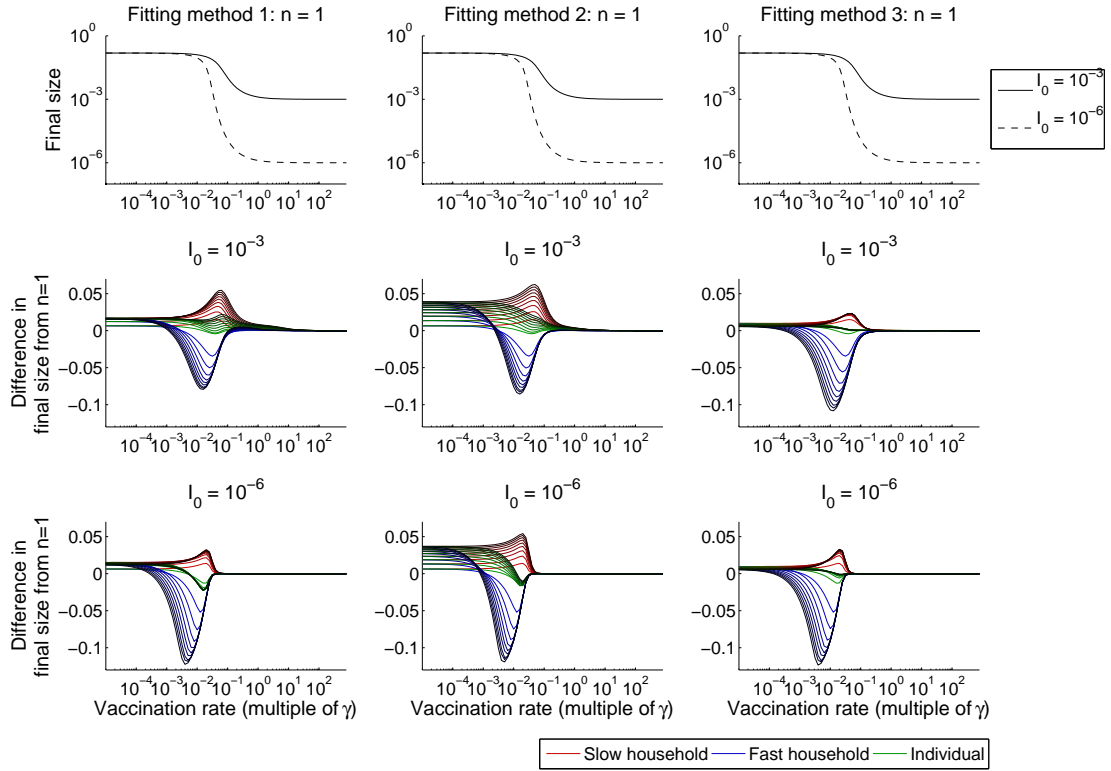


Figure 3: Effects of responsive vaccination. The top row shows the final epidemic size against the rate of distribution of vaccine for two different initial conditions in a population without household structure ( $n = 1$ ). The bottom two rows show, for each of those initial conditions, how the response changes in the presence of household structure, and for different intervention targets. Each intervention target is associated with a colour, and darker shades of that colour correspond to household sizes from  $n = 2$  to  $n = 10$ .

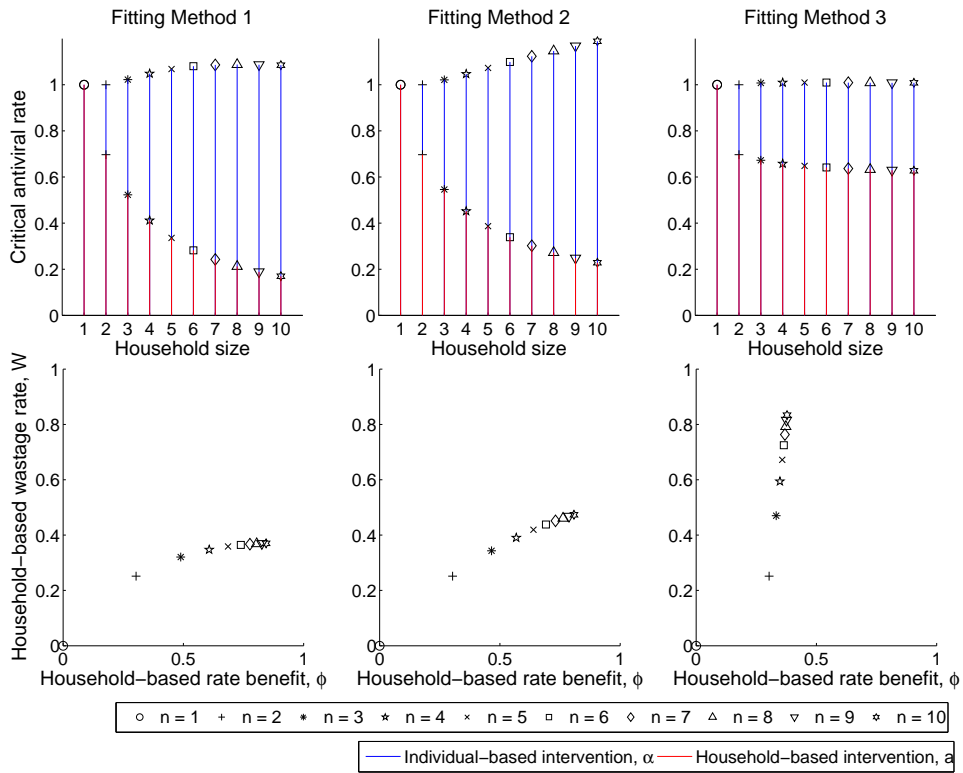


Figure 4: The effects of antiviral treatment. The first row shows the critical rate of delivery of antiviral treatment, for both targeting methods, and the second row shows the trade-off between rate benefit and wastage for household targeting.