Herd immunity acquired indirectly from interactions between the ecology of infectious diseases, demography and economics

Matthew H. Bonds1,* and Pejman Rohani2,3

1François-Xavier Bagnoud Center for Health and Human Rights, Harvard School of Public Health, Harvard University, Boston, MA 02115, USA
2Odum School of Ecology, University of Georgia, Athens, GA 30602-2202, USA
3Center for Tropical and Emerging Global Diseases, University of Georgia, Athens, GA 30602-2606, USA

Patterns of morbidity and mortality around the globe are determined by interactions between infectious diseases and systematic human socioeconomic processes. The most obvious of these patterns is that the greatest burdens of infectious diseases are found among the poor, who lack the basic resources for disease prevention and treatment. Yet, it is becoming increasingly clear that many infectious diseases are themselves causes of poverty owing to their effects on labour productivity. A particularly subtle phenomenon that receives little attention in the epidemiology literature and is especially important for poor communities is the role of the birth rate as an important direct cause of high disease burdens. Because of their high rates of transmission and life-long immunity, the persistence of many child diseases such as measles relies on high rates of reproduction as their source of susceptible individuals. Thus, there are significant direct health benefits of lower fertility rates, which are further enhanced by interactions with economic processes. Indeed, fertility, poverty and disease all interact with each other in important and predictable ways that can be built into traditional disease ecology models. We present such a model here that provides insights into the long-term effect of policy interventions. For example, because of indirect income effects, herd immunity may be acquired with lower vaccine coverage than previously thought. Reductions in the disease burden can also occur through lower fertility. Our model thus provides a disease ecology framework that is useful for the analysis of demographic transitions.

Keywords: population ecology; epidemiology; ecological modelling

1. INTRODUCTION

While most of the world has experienced unprecedented levels of sustained economic growth, over one billion people continue to suffer from the kind of extreme poverty their ancestors did many generations ago (UN Millennium Project 2005). The persistence of extreme poverty is often solely attributed to inadequate political and economic institutions, which fail to provide the critical foundations for functioning market economies (Barro 1997). But there are important correlates of global poverty that fall within the realm of the biological sciences and epidemiology, such as high prevalence of infectious diseases, low life expectancies and high rates of reproduction (World Health Organization 1998; UN Millennium Project 2005; May 2007). The most conventional explanation for these correlations is that poverty is an underlying cause of disease and mortality, and high mortality rates drive high rates of reproduction. Thus, in accordance with the demographic transition model (DTM), countries around the world can enjoy lower rates of mortality and population growth over time as a natural consequence of economic development (Caldwell 1982; Lipton 1983; Birdsall & Griffin 1988). Such theories, developed by social scientists, have applications to epidemiological dynamics because they provide predictions for systematic changes in important parameters such as rates of birth, death and transmission over time. However, the general interpretations of the direction of causality of such theories often underestimate the long-term impact of effective policy. In fact, the burden of infectious diseases is likely to be a cause of poverty and has been implicated as an underlying barrier to economic development (Strauss & Thomas 1998; Bloom & Canning 2000; Gallup & Sachs 2001; Sachs & Malaney 2002). After all, economic activity requires human resources—specifically, ‘human capital’—and therefore relies on biological processes in the form of physical labour and
cognition, which are often compromised by infectious diseases (Nokes et al. 1992; Holding & Snow 2001). Moreover, epidemiological models predict that rates of reproduction, which are notoriously high among poor populations, are among the most important direct determinants of disease burdens (McLean & Anderson 1988b; Anderson & May 1992). Thus, not only is poverty a cause of disease, mortality and high rates of reproduction, but high rates of reproduction are a cause of disease, which is a cause of poverty. An understanding of global development over the short and the long term therefore requires a broader scientific understanding of the complex interactions between infectious diseases and human reproductive and economic behaviour.

Here, we present a framework for exploring interactions between the ecology of infectious diseases, demography and economics. We focus on the role of two important variables that are simultaneously mediated by economic and biological factors: (i) the disease transmission rate and (ii) the host rate of reproduction. Our model suggests that such interactions have significant epidemiological consequences that imply greater benefits from policy interventions than standard models would predict. In particular, we show that so-called ‘herd immunity’ may be ultimately acquired with lower vaccine coverage than previously thought. As the burden of disease decreases from vaccine coverage, there can also be further long-term indirect impacts from decreases in fertility. This model thus provides a disease ecology framework for the analysis of demographic transitions.

2. THEORETICAL FRAMEWORK

The most dangerous diseases in undeveloped countries are those that have the most significant impact on mortality rates and childhood learning, and therefore ultimately on labour productivity, and are the diseases that mostly afflict children (Sachs 2005). To explore relationships between disease ecology, demography and economics, we therefore consider a childhood disease system that has been successfully modelled by compartmentalizing the host population according to its disease status: susceptible, infected and recovered (Anderson & May 1992; Grenfell & Dobson 1995). A schematic of the S–I–R model is presented in figure 1.

The formal model of the changes in densities of susceptible, S, infected, I, and recovered, R, individuals over time is represented by the following differential equations:

\[
\frac{dS}{dt} = \alpha (1 - \rho) N - \left( \mu + \left( \frac{\beta}{N} \right) I \right) S,
\]

\[
\frac{dI}{dt} = \left( \frac{\beta}{N} \right) IS - (\mu + v + \gamma) I
\]

and

\[
\frac{dR}{dt} = \alpha p N + \gamma I - \mu R.
\]

The parameter \(\alpha\) is the birth rate. The parameters \(\mu\) and \(\gamma\) are the respective rates of natural death and recovery. The frequency-dependent transmission rate is \(\beta\). The parameter \(v\) is the additional death rate caused by disease, and \(\rho\) is the proportion of individuals vaccinated at birth. The total population is represented by \(N\). Note that, while transmission rates have been successfully modelled as density dependent in developed countries that have ‘well-mixed’ populations, we assume here that poor countries, which generally have more stationary rural populations, may be better modelled with density-dependent transmission.

If the birth and death rates are equal and the pathogen is avirulent, as is broadly true for developed countries, then equations (2.1)–(2.3) also represent the dynamics of the proportions. However, if those conditions do not hold, as is typically the case for developing countries with rising populations and relatively high case-fatality rates from childhood diseases, then the dynamics of proportions of individuals are given as

\[
\frac{ds}{dt} = \alpha (1 - \rho) - (\beta - v) is - \alpha s
\]

and

\[\frac{di}{dt} = \beta si - (\gamma + v + \alpha - vi)i,\]

where \(s\) and \(i\) represent the proportions of susceptible and infected individuals. Their changes over time are derived from equations (2.1)–(2.3), where \(ds/dt = (ds/dt - s dN/dt)/N\) and \(di/dt = (di/dt - si dN/dt)/N\). The change in recovered individuals over time is simply \(1 - ds/dt - di/dt\).

This simple framework allows one to determine the equilibrium disease prevalence as a function of a handful of key parameters and is useful for policy prescriptions. For example, one powerful and important prediction is that, in order to eradicate an infectious disease, one need not vaccinate the entire population. Instead, only a proportion of the susceptible population needs to be vaccinated. This arises because of the concept of ‘herd immunity’ (Fine 1993), whereby susceptible individuals are afforded protection from infection if a sufficiently large fraction of the population is immune (Anderson & May 1979; Keeling & Rohani 2007).

The equilibrium disease prevalence that corresponds to equations (2.4) and (2.5) is

\[
i^* = \frac{\alpha(\beta - 2v) + (\beta - v)(\gamma + v) - \sqrt{(\alpha\beta + (\beta - v)(\gamma + v))^2 - 4\alpha(1 - \rho)\beta v(\beta - v)}}{2\nu(\beta - v)}.
\]
The critical vaccination threshold, above which the disease cannot persist, is given by
\[
\rho_1 = 1 - \frac{\mu + \nu + \gamma}{\beta} = 1 - \frac{1}{R_0},
\]
where \(R_0\) is the basic reproductive ratio, defined as the number of secondary infections that would result from a single infected individual in a totally susceptible population (Anderson & May 1979; Anderson et al. 1988; Keeling & Rohani 2007).

2.1. Income effect creates feedback between disease prevalence and the transmission rate

For diseases that primarily transmit through direct social or sexual contact, such as measles, tuberculosis and HIV/AIDS, it is obvious that human behaviour determines the transmission rate and therefore the basic reproductive ratio, \(R_0\). But also for vector-borne diseases such as malaria, lymphatic filariasis and dengue fever, host-to-host proximity matters because the geographical range of the mosquito vectors is limited, which is known to result in disease clustering at fine spatial scales such as at the individual household level (Gamage-Mendis et al. 1991; Carter et al. 2000; Ghebreyesus et al. 2000; Harrington et al. 2005). Such indirect host contact can even extend beyond animal vectors through what Ewald (1994) termed ‘cultural vectors’ for pathogens with free-living stages in the environment, as occurs with dysentery, cholera and hookworm, which are transmitted through faecal routes.

What all of the diseases mentioned above have in common is that they are found disproportionately among the poor, who often lack the resources to prevent such disease transmission. Sanitation measures such as proper hygiene, faecal waste management and the filtration of drinking water are all examples of economically determined human actions that lower transmission rates (Hart 1990; Gangestad & Buss 1992; Loehle 1995; Tsagkamilis 1999). But there are also more subtle influences. Poor housing conditions, such as mud walls (which allow infected mosquitoes to rest), thatched roofs, exposed sleeping conditions and crowding, are known to increase the probability of malaria infection (Gamage-Mendis et al. 1991). Lack of footwear exposes children to ground-dwelling parasitic worms. And limited access to prophylactics increases the transmission of sexually transmitted diseases such as HIV/AIDS.

But while the transmission rate is a function of per capita income, per capita income is itself ultimately a function of the ecology of infectious diseases. This is due to both (i) a direct effect of health on labour productivity and, perhaps more importantly, (ii) an indirect effect on labour productivity through the effect of childhood health on the acquisition of human capital—the training and education of the workforce. Diseases such as malaria and parasitic worms, among others, are known to directly interfere with childhood learning processes, and therefore ultimately undermine long-term economic success (Nokes et al. 1992; Holding & Snow 2001; Ezeamama et al. 2005; Fernando et al. 2006).

The infectious disease modelling literature almost always assumes that the transmission rate is determined by exogenous (i.e. fixed) parameters, which seems reasonable over short time periods. However, over the relevant economic time (i.e. intergenerational time), the transmission rate is determined by this feedback between economics and epidemiology and is thus endogenous to the system (Nokes et al. 1992; Holding & Snow 2001; Ezeamama et al. 2005; Fernando et al. 2006). We can build such feedbacks into our model system by making the transmission rate a function of income and income a function of health: \(\beta = \beta(M(i))\), where \(M\) represents per capita income, represented in figure 2. The critical vaccination threshold that accounts for feedbacks between income and transmission can therefore be represented as
\[
\rho_2 = 1 - \frac{\mu + \nu + \gamma}{\beta(M(i))} = 1 - \frac{1}{R_0(M(i))}.
\]

The difference between the classic vaccination threshold, \(\rho_1\), and the modified vaccination threshold with economic feedbacks, \(\rho_2\), depends on the functional form of the transmission rate, \(\beta(M(i))\). For all systems where decreases in the disease burden result in greater income and, therefore, lower transmission, the endogenous basic reproductive ratio, \(R_0(M(i))\), will be less than the static basic reproductive ratio, \(R_0\). Therefore, the critical vaccination threshold that accounts for economic feedbacks, \(\rho_2\), will always be less than the classic vaccination threshold, \(\rho_1\). To demonstrate this feedback, consider transmission as a simple linear negative function of income, \(\beta = \beta - \kappa M\), and income as a simple linear negative function of disease prevalence, \(M = M(1 - i)\) (an alternative functional form where transmission is an asymptotic function of income generates similar results and is presented in the appendix in the electronic supplementary material). The parameter \(\beta\) represents the maximum transmission rate in the absence of income, and \(M\) represents the maximum income in the absence of disease. This simple transmission rate can be represented as a direct function of disease
\[
\beta(i) = B + Ki,
\]
where \(B = \beta - K\), and \(K = \kappa M\).

The transmission function of equation (2.9) implies that the transmission rate is maximized in the absence of any economic resources (i.e. absolute poverty), where disease prevalence equals 1. Per capita income, and therefore the prevalence of disease, plays a critical role in determining the transmission rate.
Herd immunity acquired indirectly

M. H. Bonds and P. Rohani

544 J. R. Soc. Interface (2010)

The equilibrium disease prevalence is determined by interactions between economics and disease ecology. The black solid curve represents \( i(\beta) \) when \( \alpha = 0.06 \); the dashed curve represents \( i(\beta) \) when \( \alpha = 0.02 \); and the grey solid curve represents \( \beta(i) \). Holding other factors constant, as the disease prevalence, \( i \), rises, labour productivity and therefore \textit{per capita} income falls, and the transmission rate, \( \beta(M) \), rises. However, as the transmission rate rises, disease prevalence necessarily also rises. These feedbacks imply long-term effects from changes in parameters such as the birth rate, \( \alpha \). The two disease prevalence curves correspond to birth rates of \( 0.06 \) (solid curve) and \( 0.02 \) (dashed curve). A decrease in the birth rate from \( \alpha = 0.06 \) to \( \alpha = 0.02 \), therefore, results in a decrease in disease prevalence from \( i^*_1 \) to \( i^*_2 \), which is due to both a direct decrease in the in-flow of susceptibles and an indirect increase in \textit{per capita} income resulting in lower transmission: \( \gamma = 0.02; \mu = 0.014; \nu = 0.04; K = 1; B = 0.3 \).

Because the transmission rate is a function of income, which is determined by the disease burden, and the disease burden is determined by the transmission rate, the long-term equilibrium disease prevalence and \textit{per capita} income are determined simultaneously, as represented in figure 3. The effect of vaccination on the disease prevalence for the traditional model compared with the present model is presented in figure 4.

### 2.2. Effects of host fertility on infectious diseases

Understanding that \textit{per capita} income and the prevalence of infectious diseases are determined simultaneously over long epidemiological/economic time periods is necessary for a better understanding of the relationship between economic development and global health, and therefore of the long-term consequences of sustained policy interventions that affect key epidemiological parameters. One important parameter that is highly correlated with poverty is the birth rate. Many of the world’s most important childhood diseases, such as measles, are characterized by high disease transmissibility coupled with the acquisition of host immunity, resulting in low mean ages of infection. The diseases are therefore sustained by a constant flow of immunologically naive, susceptible children. But children do not merely serve as new hosts for the pathogens. They also serve as new conduits for transmission. As a result, the birth of a child in the poorest parts of the world represents not only a new infection opportunity for a disease, but also an increase in the probability of infection for the rest of the susceptible host population. Thus, epidemiological theory predicts that a reduction in the birth rate can significantly lower the prevalence of childhood diseases (Anderson & May 1992; Anderson et al. 1988).

Figure 5 demonstrates the simultaneous effect of vaccination and birth rate on disease prevalence for both the traditional model (figure 5a) and the model that accounts for indirect economic effects (figure 5b). Figure 6 demonstrates the difference between the two models, \( i^*(\beta) - i^*(\beta(M(i))) \), over the range of vaccination and birth rates. Notice that not only is the critical vaccination coverage, \( \rho_{c\rightarrow 0} \), always lower for the model with economic feedbacks, but there are significant differences in the predicted prevalence of disease for vaccination rates above the critical value. The size of the difference depends on birth rates. The most significant differences between the models are found when birth rates are high and when vaccination coverage is high enough to have a significant impact on the disease burden. These conditions are most common in poor countries.

While focusing on such interlinkages between socioeconomic and biological processes may seem analytically complicated, these models have several important implications. They suggest, for example, that \textit{any} intervention that targets either economic
productivity, fertility or disease prevalence will indirectly result in an improvement in the related health and economic conditions, and, therefore, will have a greater impact on the targeted outcome than would be predicted by conventional epidemiological forecasts. The most effective way to improve health and economic development may therefore be to capitalize on synergistic effects via multiple interventions simultaneously.

3. DISCUSSION

The biggest killers of human beings in the undeveloped world are infectious diseases. From theoretical developments based on the classic ‘S–I–R’ model framework, substantial progress has been made in understanding the ecology and evolution of such diseases, which have appropriately relied on the principles of predator–prey relationships (Lotka 1925; Volterra 1926; Anderson & May 1979, 1992). Important policy implications have emerged. For example, we have learned that disease eradication does not necessarily require vaccination of an entire population. Rather, herd immunity can be acquired by vaccinating only a proportion of the population (equation (2.7)). While much analysis has successfully focused on diseases of childhood (Rohani et al. 1999, 2003; Grenfell et al. 2001), because of data availability these efforts have focused on dynamics in developed countries, where the populations are relatively stable and the mortality effect of the diseases is near zero. Key factors for these systems include the transmission rate and the birth rate, which are almost always considered to be exogenous parameters.

Infectious diseases, however, continue to be most significant in poor countries, where birth rates and transmission rates are systematically high (McLean & Anderson 1988a; Broutin et al. 2005). Indeed, poverty, fertility and disease are not only signatures of each other over different parts of the globe today, they are also considered highly correlated over time and, as described by the DTM, are associated with specific stages of economic development (Caldwell et al. 2006). Although theoretical explanations of demographic transitions have existed for many decades, they have focused on socioeconomics; specifically, on economic growth and reproductive behaviour. An important principle that has been entirely absent from demographic transition theories that further explains correlations between fertility and mortality is that, through the ecology of infectious diseases, fertility causes high disease burdens and therefore mortality.

The model presented here is not a specific attempt to explain demographic transitions, per se, but to explore how the basic epidemiological theory can be synthesized with demography to explore critical processes that systematically affect human welfare and economic development. It is during youth, when immunity is naive and susceptibility is high, that children must acquire the training and skills necessary for economic success later in life. In addition to the naive immunity,
pathogens can therefore exploit indirect advantages of targeting children by debilitating their human hosts when they are young, which ultimately, through poverty, undermines their ability to protect themselves and their own children later in life from the assault of the same infectious agents. Over the longer term, the basic reproductive ratio, $R_0$, is therefore determined simultaneously by both the biology of disease transmission and the economics of host vulnerability.

Despite the importance of understanding the ecological principles that govern infectious diseases dynamics and their impacts for economic development, infectious disease ecologists have yet to build formal models that meaningfully contribute to our understanding of these processes. Perhaps the best attempt at integrating relevant theory was developed by Delfino & Simmons (2005), who explored dynamical consequences of tuberculosis for different wealth categories that were, in turn, determined by the disease burden. However, Delfino & Simmons (2005) stopped short of leveraging important principles of disease ecology (such as the calculation of $R_0$) to shed light on policy implications (such as the critical vaccination threshold).

The model here, designed to help build a framework for exploring linkages between socioeconomic behaviour and the ecology of infectious diseases, does have important policy implications. For example, after accounting for the effects of lower disease burdens that would be predicted by social scientists, such as greater economic productivity and lower fertility, sustained vaccination programmes could achieve herd immunity by covering a lower proportion of the population than would be deemed necessary by standard epidemiological models. The long-term critical vaccination threshold is therefore just as dependent on the economic and reproductive responses to vaccines as it is on short-term calculations of $R_0$. Fortunately, the last 30 years have experienced massive increases in global vaccination coverage for the major child diseases. Global measles vaccination coverage, for example, has risen from around 13 per cent in 1980 to 84 per cent today, with coverage being the lowest in the poorest regions of the world such as sub-Saharan Africa (World Health Organization 2006). The rate at which $R_0$ and thereby the critical vaccination threshold, falls will depend on the rate at which fertility decreases and economic growth rises, which would be expected to take several generations. Indeed, while fertility rates remain high in certain areas, average fertility is already falling throughout the globe, if more slowly in areas of extreme poverty.

While our general approach for integrating these theories should be relevant for many long-term disease-economic systems, the specific details and corresponding policy implications would depend on the specific system of interest. Our model is simple in that it represents a single disease, focusing on interactions with only a couple of key parameters. In reality, many diseases affect socioeconomic behaviour in the developing world, all with slightly different transmission dynamics and different channels of impacts. An important extension of this framework would therefore be to consider other relevant diseases, such as malaria, TB, AIDS, respiratory infections and parasitic worms, each of which is not only independently significant but has been shown to interact with the others. The specific functional relationship between disease burden and the relevant parameters, such as rates of birth, transmission, recovery and virulence, may all depend on a combination of ecological and sociodemographic factors, all of which may have qualitative implications, and which justify further exploration.

Underlying the intense and broad upsurge of interest in global health and economic development are important questions. Why do people from some parts of the world enjoy continued exponential economic growth, while others, such as those in sub-Saharan Africa, suffer from the kind of extreme poverty their ancestors did thousands of years ago? While some very important explanations have come forth that explicitly link economic development to disease ecology (Diamond 1997; Sachs 2005), we continue to lack formal integrated theories. Fortunately, the structure of economic and ecological modelling is primed for such integration. Our simple framework here presents an example of such integration, and hopefully constitutes more than an epidemiological model of poverty, but a step towards more general integrated theories of ecology and economics.

The authors would like to thank Donald Keenan and Lord Robert May for comments on the manuscript. We especially thank the reviewers from the Journal of the Royal Society Interface for important and useful comments. This project was supported by an ARCS Fellowship and an Earth Institute Fellowship to M.H.B., and a National Science Foundation, Human and Social Dynamics Grant to P.R.

**REFERENCES**


